

ORIGINAL ARTICLE

Murine candidate bleomycin induced pulmonary fibrosis susceptibility genes identified by gene expression and sequence analysis of linkage regions

C K Haston, T G Tomko, N Godin, L Kerckhoff, M T Hallett

J Med Genet 2005;42:464–473. doi: 10.1136/jmg.2004.027938

See end of article for authors' affiliations

Correspondence to:
Dr Christina Haston,
Department of Medicine,
McGill University,
Meakins-Christie
Laboratories, 3626 rue St.
Urbain, Montreal, QC,
Canada H2X 2P2;
Christina.haston@mcgill.
ca

Revised version received
21 December 2004
Accepted for publication
23 December 2004

Background: Pulmonary fibrosis is a complex disease for which the predisposing genetic variants remain unknown. In a prior study, susceptibility to bleomycin induced pulmonary fibrosis was mapped to loci *Blmpf1* and *Blmpf2* on chromosomes 17 and 11, respectively, in a C57BL/6J (B6, susceptible) and C3H/KAM (C3H, resistant) mouse cross.

Methods: Herein, the genetic basis of bleomycin induced pulmonary fibrosis was investigated in an approach combining gene expression and sequencing data with previously mapped linkage intervals.

Results: In this study, gene expression analysis with microarrays revealed 1892 genes or ESTs (expressed sequence tags) to be differentially expressed between bleomycin treated B6 and C3H mice and 67 of these genetic elements map to *Blmpf1* or *Blmpf2*. This group included genes involved in an oxidative stress response, in apoptosis, and in immune regulation. A comparison of the B6 and C3H sequence, for *Blmpf1* and *Blmpf2*, made using the NCBI database and available C3H sequence, revealed approximately 35% of the genes in these regions contain non-synonymous coding sequence changes. An assessment of genotype/phenotype correlation among other inbred strains revealed 36% of these B6/C3H sequence variations predict for the known bleomycin induced fibrosis susceptibility of the DBA (susceptible) and A/J (resistant) mouse strains.

Conclusions: Combining genomics approaches of differential gene expression and sequence variation potentially identifies approximately 5% the linked genes as fibrosis susceptibility candidate genes in this mouse cross.

Pulmonary fibrosis is a disease in which the lung response to known (such as environmental toxins or radiation therapy) or unknown injuries produces a functionally compromising pathology. This devastating disease has an estimated incidence of 1 case per 10 000 per year and associated mortality is 50–70% at 5 years post diagnosis.¹ The disease is characterised by cellular proliferation and progressive accumulation of extracellular matrix constituents resulting in remodelling of the lung interstitium.¹ The mechanisms through which this functionally compromising pathology develops are unknown, but inflammatory,^{2–4} apoptotic,¹ and matrix metalloproteinase/tissue inhibitor of metalloproteinase imbalance^{2–5} pathways have been implicated. Based on reports of familial pulmonary fibrosis, the disease is thought to have a genetic basis but the specific genes involved have not been identified.^{1–6}

To circumvent the clinical limitations of environmental variability, genetic heterogeneity, and small sample sizes, investigations of inbred mouse strains which differ in their susceptibility to the fibrotic phenotype^{7–8} can be used to identify fibrosis susceptibility genes. Specifically, the anti-neoplastic drug bleomycin, which when used clinically leads to the development of fibrosis in approximately 5% of the exposed clinical population,⁹ when delivered through a 7 day subcutaneous dose has also been found to produce a fibrotic phenotype in mice which resembles idiopathic pulmonary fibrosis.¹⁰

We have previously mapped two loci of murine strain difference in propensity to develop fibrosis after bleomycin treatment in a C57BL/6J (B6, susceptible) and C3H (resistant) mouse model.¹¹ The influence of the loci, named *Blmpf1* (bleomycin induced pulmonary fibrosis 1) on chromosome 17 (LOD = 18.4) and *Blmpf2* on chromosome 11 (LOD = 5.6) on

the fibrotic phenotype, and the interaction of the two loci, were confirmed in additional studies of congenic and chromosome substitution mice.¹¹

Consistent with the initial mapping of many complex traits, the *Blmpf1* and *Blmpf2* intervals are too large (approximately 13 and 26 Mb, respectively) to permit a practical assessment of individual candidate genes in the search for causative genetic variation. Thus, our approach to identifying candidate genes of fibrosis susceptibility has been to combine the results from additional genomic experiments with the mapping data. Specifically, gene expression analysis with microarrays, parental strain DNA sequence analysis, and genotype/phenotype correlation in additional mouse strains, were used to produce a set of candidate genes involved in the pathogenesis of experimental fibrosis. The resources available for this study included the genomic sequence of B6 mice¹² and the partial genomic sequence of the C3H strain,¹³ and the available genome sequence (Celera) and documented bleomycin induced fibrosis susceptibility of inbred strains DBA/2 (susceptible),¹⁴ A/J (resistant),⁷ and 129 (susceptible).¹⁵

In this study, we present evidence for a set of candidate genes underlying the susceptibility loci *Blmpf1* and *Blmpf2*. The fibrosis genes were isolated from among the mapped positional candidates, by identifying the subset of positional candidate genes which are differentially expressed between B6 and C3H mice in the bleomycin treated lung, and for which there is a sequence variation between the strains. These candidate genes were also assessed for genotypic correlation in mouse strains of known fibrosis susceptibility phenotype.

Abbreviations: ESTs, expressed sequence tags

METHODS

Mice

Mice of the C57BL/6J (B6) and C3H/HeJ (C3H) strains were purchased from Jackson Laboratories (Bar Harbor, ME) and housed in the animal facility of the Meakins-Christie Laboratories.

Bleomycin treatment

Mice were treated at 8 weeks of age. Lung damage was elicited by administering bleomycin through osmotic minipumps implanted subcutaneously, as described previously.¹⁶ Male mice received 100 units of bleomycin per kg body weight (~2.5 units per mouse) and female mice received 125 units per kg body weight. Male and female mice were treated in separate studies due to the higher drug dose required for female mice and due to the sex specificity of the previously mapped loci. The mice were sacrificed ($n = 4-8$ mice per group) when moribund or at 3 or 6 weeks after treatment. Twenty untreated mice, five males and five females of each strain, were used as controls and were sacrificed at the 6 week time point.

Histology and fibrosis scoring

At autopsy, the lungs were removed and the single left lobe of each mouse was perfused with 10% neutral buffered formalin and submitted for histological processing. Each left lung was stained with Masson's trichrome for identification of the site(s) of collagen deposition in the lung. The area of the fibrosing phenotype for each mouse was quantified with image analysis of histological sections as described previously.¹⁶ Specifically, the area of fibrosis in the left lung lobe was determined from a user drawn region surrounding the fibrosis and compared to the area of the entire lobe to yield the percent of pulmonary fibrosis for individual mice.

Gene expression

Following sacrifice, the right lung of each mouse was immediately homogenised in 2 ml of Trizol reagent and placed in dry ice. The homogenates were stored at -85°C until RNA isolation. Total lung RNA was extracted according to the manufacturer's instructions (Sigma, St Louis, MO). The RNA from the right lungs of four or five mice from each group, defined by sex, strain, treatment, and time point, was pooled, as in Perkowski *et al.*,¹⁷ to minimise biological variation in gene expression within a group. One sample of pooled RNA for each group was processed through the RNeasy column (Qiagen, Valencia, CA) and submitted for hybridisation. RNA quality was assessed and confirmed prior to and following pooling using the Agilent Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). The experiment was performed with one chip per mouse group, represented by its pooled RNA.

Microarray hybridisation was performed by the Affymetrix Gene Chip Core facility at the McGill University and Genome Quebec Innovation Centre. The probe synthesis, hybridisation, and washing protocols used were as described in Novak *et al.*¹⁸ and followed the standardised Affymetrix protocol. The starting material was 10 μl of total RNA. After hybridisation to the Murine MOE430A GeneChip (Affymetrix, Santa Clara, CA), the gene chips were automatically washed and stained with streptavidin-phycoerythrin by using a fluidics system. The chips were scanned with a GeneArray Scanner (Agilent Technologies). The resultant gene expression profiles were then extracted and viewed using Microarray Suite 5.0 (Affymetrix). The MOE430A GeneChip arrays contain 22 690 probe sets derived from sequence clusters contained in Build 107 (UniGene, June 2002) which represent approximately 14 000 functionally annotated genes and a set of expressed sequence tags (ESTs).

Microarray data analysis

Bioconductor version 1.4 (see <http://www.bioconductor.org/>) routines within the R (version 1.90) statistical language¹⁹ were used for quality control, normalisation, and differential expression. In particular, the quality of the raw microarray data was assessed by inspecting similarities between the intensity distribution and RNA digestion plot for each array. Normalisation was performed using the robust probe level model.²⁰ Detection of differential expression was performed using the LIMMA package^{21,22} with a p value <0.05 . Lists of significantly differentially expressed genes were generated intra-strain (control *v* bleomycin exposure) and between strains. Hierarchical clustering was performed via Cluster version 3.0²³ using the average linkage and Euclidean distance options. Results were visualised using TreeView version 1.08 (<http://rana.lbl.gov/>). The detection of significantly over represented Gene Ontology categories was performed using the FatiGO tool.²⁴ The test of statistical significance considers the number of differentially expressed genes found in this category compared to the total number of genes in the category represented on the chip. Raw and normalised expression data are available at the NCBI GEO repository.

Quantitative real time PCR

To generate the cDNA for real time PCR, 4–5 μg of total RNA from the right mouse lung was reversely transcribed with oligo(dT)_{12–18}Primer using Superscript II RNase H⁻ Reverse Transcriptase (Invitrogen, Carlsbad, CA) in a 20 μl total volume. Reverse transcription was performed at 65°C for 5 min, followed by the activation of SuperScript II RT at 42°C for 50 min. The reaction was inactivated by heating to 70°C for 15 min.

Sequence specific primer sets, designed using Primer 3 or taken from the RT-PCR database, were used to carry out quantitative real time PCR. The primer sequences were confirmed, by BLAST comparison to the mouse genome sequence in the Ensembl database, to span an intron, and to be specific for the gene being assayed. Probe and primer sequences are as follows (5'→3'):

- glutamate receptor, ionotropic, AMPA1 (alpha 1) (GenBank accession number NM_008165, PrimerBank ID 34328128a1), forward: CACTGGGTGATGCTCCTCG, reverse: CCTCACTGGGTAGTCCATACAC
- pituitary tumour transforming 1 interacting protein (GenBank accession number NM_145925, PrimerBank ID 22122339a1), forward: GTCCGCCCTGAGAAATCCAG, reverse: CTCGCCCTTGTCGTACCAC
- glyoxalase 1 (GenBank accession number NM_025374, PrimerBank ID 31981282a1), forward: GATTGTGTCACA TTGGGATTGC, reverse: TCCTTTCATTTCCCGTCATCAG
- complement component 4 (within H-2S) (GenBank accession number NM_009780, PrimerBank ID 6753226a3), forward: TGCCAATGAAGACTACGAAGAC, reverse: TGCCATTTCGCCAGATACACA
- spinocerebellar ataxia 10 (GenBank accession number NM_016843), forward: AAGCTAGTGGGTGAGGAGCA, reverse: ATCGTAACCAGGGCTTCTT
- proteasome subunit α type 3 (GenBank accession number NM_011184), forward: GCAGTTTCATGTTGGGGTCT, reverse: CGGCAAGTCATTTCCTTCAT,

For real time PCR, a LightCycler thermocycler from Roche (Indianapolis, IN) was used. Each reaction in a 32 well carousel contained 1 μl of cDNA template and 10 μl of QuantiTect SYBR Green PCR Kit (Qiagen). The PCR reaction followed a heating temperature of 95°C , an annealing temperature of 55°C , and an elongation temperature of 72°C

for 45 or 50 cycles. Each experiment included a non-reverse transcribed control to assay for genomic contamination. No amplification was observed in this type of reaction or in reactions in which no cDNA template was used.

Relative gene expression data analysis was carried out with the standard curve method.²⁵ In this analysis, the expression level of each of the target genes was determined relative to that of a reference gene. The reference gene was chosen from the set of genes whose expression varied from 0.9 to 1.1 relative units across the 13 arrays. The expression of the putative reference gene was then investigated, with RT-PCR, in each of the sample types used in this study. The ratio of the sample fluorescence to that of a standard curve generated using a serial dilution of the template (a PCR fragment cloned into pCR 2.1-TOPO vector) was determined for each of putative reference genes *Sca10* and *PSMA3*. The expression of *Sca10* was found to vary from 0.85 to 1.15, relative to *PSMA3*, across eight treatment groups and was used as the reference gene for the relative expression of the target genes.

The fluorescence data for each target gene were normalised to its own standard curve (generated through serial dilution of the cloned PCR product), then normalised to the expression level of the reference gene, and finally, expressed as a ratio of the calibrator sample which was included in every run. The expression of each target gene was taken as the average of the expression levels of two mice from the same treatment group, and each of these samples was run in quadruplicate.

B6/C3H sequence comparison

An interrogation of Ensembl version 21.22b.1 (http://www.ensembl.org/Mus_musculus/) was made to ascertain the number of genes mapping to the linkage regions *Blmpf1* and 2. The Ensembl positions of the markers flanking *Blmpf1* (D17Mit198/D17rua1) and those flanking *Blmpf2* (D11Mit136/D11Mit5) were determined and the sequence data therein considered. To ascertain which genes in these regions are polymorphic between B6 and C3H mice, the NCBI database (dbSNP build 121, <http://www.ncbi.nlm.nih.gov/SNP/MouseSNP.cgi>) was queried for B6/C3H SNPs and MIT produced C3H sequence was compared to Ensembl derived B6 sequence. For the SNP analysis, the database was queried for B6/C3H polymorphisms on chromosomes 11 and 17 and re-queried for SNPs on these chromosomes which are non-synonymous. The genes containing the SNPs of the latter query were assessed for map position to the *Blmpf1* and 2 linkage regions.

Sequence variations were also identified by comparing (by BLASTN) the Ensembl derived B6 sequence for each gene of *Blmpf1* and 2, to the MIT provided C3H sequence (http://www.broad.mit.edu/ftp/pub/mouse_contigs/Mar10_02/).¹² Each of the B6 genes and 1 kbp of the flanking 5' and 3' sequence were compared to the database of C3H. The C3H contigs with an E value of zero and a match score of 800 were further evaluated. For each of the genes in *Blmpf1* and 2 it was determined if the inbred strains (B6 and C3H) differ in their sequence, and, for the coding regions, if that sequence change was synonymous (encoded the same amino acid) or non-synonymous (encoded different amino acids). This analysis was completed by translating the B6 sequence with and without the exonic mismatch (deletion/insertion/nucleotide change) and comparing the resultant peptides.

Multi strain sequence comparison

The markers flanking each of *Blmpf1* and 2 (see above) were located in the Celera (Rockville, MD) mouse genome database (<http://www.celera.com/>, CDS 13h release) and the number of genes (excluding pseudogenes) mapping to each region was determined. Using the DNA positions of the

flanking markers, the Celera Mouse SNP reference database (v3.6) was queried for SNPs within each linkage region. These SNP data were then filtered to uncover the set of SNPs for which B6, DBA/2, and 129 mice (all phenotypically similar in bleomycin response) carry the same allele and for which A/J mice (phenotypically distinct) have a different allele. Since it is not known which of the substrains of mouse strain 129, sequenced by Celera, was determined to develop fibrosis after bleomycin exposure,¹⁵ this analysis was completed with and without the 129 strain. In both analyses, these data were further filtered to exclude SNPs appearing in the intronic region or identified as synonymous, as has been done by others.^{26, 27}

RESULTS

Histological phenotype

To measure the level of gene expression in the lung and to histologically confirm the strain variation in fibrosis phenotype, mice of the B6 and C3H strains were treated with bleomycin and sacrificed at different time points after treatment. In fig 1 the average fibrosis percent of the lung is shown in mice sacrificed 3 or 6 weeks after bleomycin treatment, and in mice sacrificed due to respiratory distress. The mice sacrificed when sick had lost more than 20% of body weight and were sacrificed from 21 to 36 days after treatment. B6 female mice developed pulmonary fibrosis defined by alveolar atelectatic regions with collagen deposition, of approximately 8% of the lung (fig 1) in response to bleomycin treatment. This fibrosis response was evident, and of similar magnitude, in B6 female mice at 3 and 6 weeks after treatment and in the mice sacrificed due to respiratory distress. The fibrotic lung response of B6 male mice to bleomycin was similar to that of the B6 female mice at 3 weeks and when the mice were in distress ($p = 0.10$ and 0.15, respectively). B6 male mice are, however, more sensitive to bleomycin as this fibrosis level was in response to a lower dose of the drug. No bleomycin treated B6 male mice survived to the 6 week time point.

In C3H mice the response to bleomycin was a minimal lung disease of limited fibrosis foci. As shown in fig 1, the C3H strain mice, males and females, developed fibrosis foci involving approximately 0.4% of the lung at 3 and 6 weeks

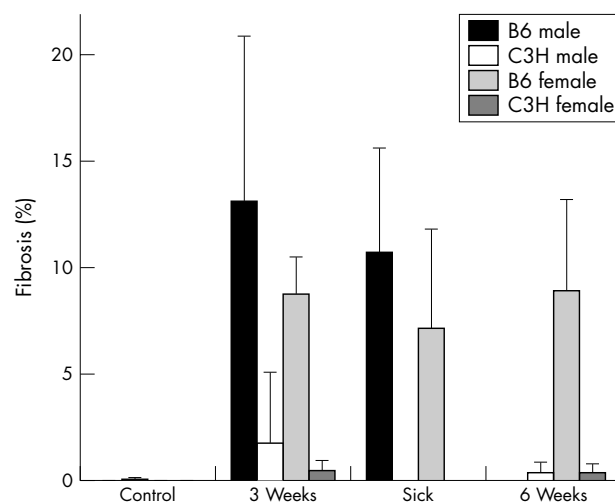


Figure 1 Strain difference in lung response to bleomycin. Mean percent of fibrosis, with the standard deviation, in a histological section of the lung following a dose of bleomycin by strain, sex, and time of sacrifice. "Sick" indicates mice sacrificed when in respiratory distress. No C3H mice were identified to be in respiratory distress. No B6 male mice survived to 6 weeks after treatment.

Table 1 Summary of B6/C3H sequence analyses for the linkage regions

	<i>Blmpf1</i>	<i>Blmpf2</i>	Total
Region (position in cM)	D17Mit198/D17rta1 (16.0–22.8)	D11Mit136/5 (19.0–37.0)	
Base pairs	27 292 264–40 157 535; 12.9 Mb	40 853 984–66 758 852; 25.9 Mb	38.8 Mb
Number of genes	328	351	679
Number of genes with $\geq 90\%$ B6/C3H sequence comparison (and non-synonymous coding region sequence variation)	133 (41%)	155 (44%)	288 (42%)
Number of genes with 50–89% B6/C3H sequence comparison (and non-synonymous coding region sequence variation)	50 (15%)	51 (15%)	101 (15%)
Number of genes with 50–89% B6/C3H sequence comparison (and non-synonymous coding region sequence variation)	188 (57%)	189 (54%)	377 (56%)
Number of genes with 50–89% B6/C3H sequence comparison (and non-synonymous coding region sequence variation)	46 (14%)	38 (11%)	84 (12%)

after drug exposure. This response was not different between male and female mice ($p = 0.11$). No C3H mice were identified to be in respiratory distress. The lung response of B6 female mice to bleomycin was greater than that of C3H female mice at both 3 and 6 weeks ($p = 9.3 \times 10^{-6}$, $p = 6.6 \times 10^{-5}$, respectively; fig 1). The strain difference in the response to bleomycin is evident in the male mice as the mean fibrosis score for bleomycin treated B6 males at 3 weeks post exposure was greater than that for C3H males ($p = 8.8 \times 10^{-3}$).

B6 v C3H sequence comparison

To identify the sequence variation between B6 and C3H mice for the genomic regions *Blmpf1* (12.9 Mb) and *Blmpf2* (25.9 Mb) (table 1), the NCBI database was queried for B6/C3H SNPs. With filtering by strain (B6 v C3H) and chromosome only, this database returned 1278 SNPs on chromosome 17 and 2588 on chromosome 11. Further filtering of this query produced 31 and 30 non-synonymous SNPs on the two chromosomes. An analysis of the genes in which the SNPs were contained compared to the *Blmpf1* and 2 gene lists revealed approximately 1.6% (11 of 679) of the genes contained B6/C3H sequence variation(s) which would be predicted to alter the protein product (coding region non-synonymous polymorphism).

To augment this analysis, a second B6/C3H sequence comparison was completed using the B6 sequence data of Ensembl and the partial C3H draft sequence data of MIT.¹³ For 42% of the linkage region genes, a C3H sequence match ($\geq 90\%$ of gene length) was identified. Of these genes ($n = 288$), a non-synonymous coding region sequence variation (SNP or insertion/deletion) was identified in 35%. A further 84 genes (of 377, 22%), for which a C3H sequence encompassing 50–89% of gene length was identified, contained a coding non-synonymous sequence variation in the portion of the gene evaluated. Based on this analysis it is conservatively estimated that the fraction of genes in *Blmpf1* and 2 that contain B6/C3H sequence variations to alter the encoded protein is 35%.

Multi strain sequence comparison

As the analysis of SNPs in the B6 genome compared to the C3H is limited by the lack of a complete genetic sequence of the C3H strain, a second SNP analysis was completed using the known response to bleomycin of the four strains for which sequence is available in the Celera database. Localisation of the two flanking markers of each of *Blmpf1* and 2 on the Celera map revealed both *Blmpf1* and 2 to be larger on the Celera map than on the Ensembl map (regions of ~20 and 29.1 Mb, respectively, in Celera). Scanning of the Celera database for SNPs in the linkage regions identified a number of SNPs consistent with the rate of one SNP/kbp of

Table 2 Summary of SNPs found in Celera Mouse SNP reference database (v3.6)

	<i>Blmpf1</i>	<i>Blmpf2</i>
Number of genes	380	352
Total SNPs (B6 = DBA/2) \neq A/J	23719	26755
With intronic/synonymous SNPs	4538	5229
Without intronic/synonymous SNPs (B6 = DBA/2 = 129) \neq A/J	148	154
With intronic/synonymous SNPs	754	528
Without intronic/synonymous SNPs	28	17

sequence which has been reported for the mouse genome¹² (table 2).

The linkage region SNP data were filtered to identify candidate genes where the sequence of fibrosis prone strains (B6, DBA/2) was identical and distinct from that of a fibrosis resistant strain (A/J). After removing intronic or synonymous SNPs from the analysis, 148 SNPs encompassing 52 genes in *Blmpf1* and 154 SNPs in 24 genes of the sequence of *Blmpf2*, were revealed. To integrate this sequence variation with the previous analysis, the subset of the 76 SNP containing genes from the Celera database, for which the B6/C3H sequence comparison was available, was compiled ($n = 28$). Ten (36%) of the 28 genes were identified to contain SNPs where the allele in B6 mice was the same as in DBA/2 mice but distinct from the A/J strain and where the B6 allele differed from the C3H allele.

When the data from the Celera database were filtered to include the fibrosis susceptible 129 strain, 28 SNPs encompassing 13 genes in *Blmpf1* and 17 SNPs in six genes of the sequence of *Blmpf2* were identified. Of these genes, two in *Blmpf1* (cytochrome p450 4f14, histocompatibility-2 M1) and one in *Blmpf2* (retinoic acid induced-1) were also identified as containing a B6/C3H sequence variation.

Gene expression data

Affymetrix GeneChip microarrays were used to determine which genes are differentially expressed in the lung tissue of B6 mice compared to C3H mice. In this experiment, the data of each array represent the gene expression in the lungs of a group of mice of the same sex, strain, bleomycin dose, and time point. The expression data for the 13 chips used in this experiment are given in fig 2. As shown in this figure, the data cluster into three main groups: genes highly expressed in B6 mice (treated or control); genes highly expressed in C3H mice (treated or control); and genes over or under expressed in B6 mice treated with bleomycin relative to both C3H mice and B6 control mice. The similarity in gene expression within each strain, evident in fig 2, was analysed

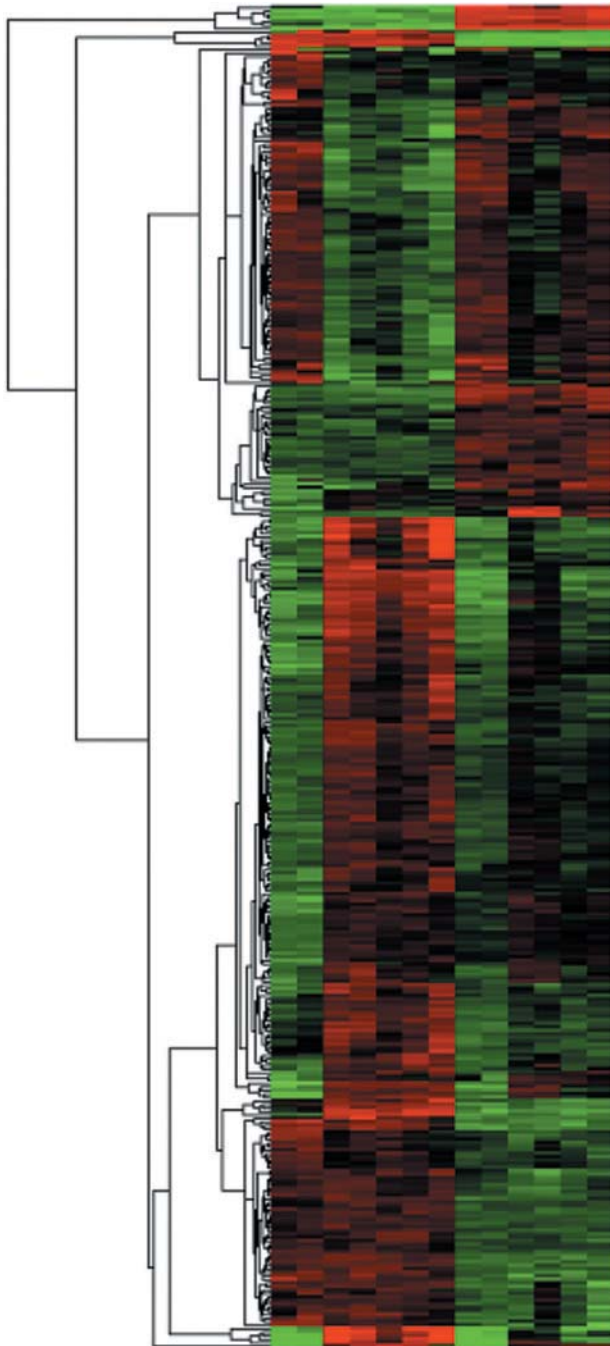


Figure 2 Murine lung gene expression after bleomycin treatment. Each column represents one group of mice defined by strain, sex, bleomycin treatment, and time post treatment and each row is a separate gene. Red indicates over expression of a gene and green is under expression ($p < 0.05$, fold > 2). Groups (left to right) are B6 male control, B6 female control, B6 male 3 weeks (post bleomycin), B6 female 3 weeks, B6 female 6 weeks, B6 male in respiratory distress, B6 female in respiratory distress, C3H male control, C3H female control, C3H male 3 weeks, C3H female 3 weeks, C3H male 6 weeks, and C3H female 6 weeks.

to determine if the gene expression profiles differed between male and female mice within a strain, or between mice evaluated at different times after bleomycin treatment.

To determine if pulmonary gene expression differed between male and female mice, we compared the gene expression profile of B6 male control mice to that of B6 female controls, and the profiles of B6 male mice treated with bleomycin were compared to those of bleomycin treated B6

female mice. The same comparisons were made of the C3H derived gene expression data. Two genes (DEAD box polypeptide 3, Y linked and eukaryotic translation initiation factor 2, subunit 3, structural gene) were found to be differentially expressed (> 2 fold difference in expression, $p < 0.05$) between male and female mice and this difference was evident in all the comparisons (data not shown). As only two Y linked genes were identified to differ in expression between male and female mice, subsequent analyses were completed with each time point comprised of data from male and female mice.

In a similar analysis we determined if the gene expression profile of C3H mice at 3 weeks post bleomycin differed from the profile at 6 weeks. This evaluation was also completed for B6 mice, and, for this strain, a comparison to the profile generated from mice identified to be in respiratory distress was also completed. In the B6 mice two genes (chemokine (C-X-C motif) ligand 10, and deoxycytidine kinase) were found to be more highly expressed at 3 weeks than at 6 weeks or in mice in distress while in C3H mice one gene (lipocalin 2) was more highly expressed at 3 weeks than at 6 weeks (> 2 fold difference in expression, $p < 0.05$; data not shown). Based on this analysis the gene expression profiles at 3 and 6 weeks post bleomycin and from mice sacrificed due to distress were determined to be similar and these datasets were combined for further analyses.

Strain specific gene expression response to bleomycin

To determine the set of genes involved in the pulmonary response of a B6 mouse to bleomycin, the data from five arrays (female 3 weeks, female 6 weeks, female sick; male 3 weeks, male sick) were compared to the data from two control arrays (male control, female control). A total of 766 genes were determined to be significantly over expressed in the lungs of treated mice compared to untreated controls and 635 genes were relatively under expressed, as shown in supplementary table 1 ($p < 0.05$) (supplementary table 1 is available at <http://www.jmedgenet.com/supplemental>). In addition, 367 ESTs were differentially expressed ($p < 0.05$) in the lungs of bleomycin treated B6 mice compared to controls (supplementary table 1).

By Gene Ontology analysis, the biological processes most affected in the B6 bleomycin treated mouse lung were cell communication ($p = 0.011$) and response to stimulus ($p = 0.018$). The molecular functions most altered were protein ($p = 0.0003$), nucleic acid ($p = 0.0073$), and glycosaminoglycan ($p = 0.03$) binding; extracellular matrix ($p = 0.013$) and muscle ($p = 0.018$) constituents; and small protein conjugating enzyme ($p = 0.019$), lipid transporter ($p = 0.03$), and GTPase regulator ($p = 0.045$) activity. As shown in table 3, the most significantly differentially expressed genes (ranked by p value) included those of the extracellular matrix such as fibronectin 1 and collagen types 1 α 2 and V α 2, and genes involved in extracellular matrix homeostasis as indicated by cathepsin S and B over expression in lungs of bleomycin treated mice compared to controls. Also differentially expressed were genes involved in immunity, such as cd68 antigen, bone marrow stromal cell antigen 1, and an IgG receptor. Each of these genes is an example of the types of genes differentially expressed in the B6 lung after bleomycin exposure. Additional types of genes differentially expressed in the B6 lung included cytokines, integrins, growth factors and chemokines, and their receptors; oxidative stress response genes; and genes encoding cell signalling proteins, ion channels, and cytochrome p450 proteins. The gene expression data are consistent with the inflammatory and fibrotic phenotype evident in the lungs of bleomycin treated B6 mice.

Table 3 Genes most significantly differentially expressed in the lungs of bleomycin treated B6 mice relative to non-treated B6 mice

Symbol	Name	Fold (treated/control)	p Value	UniGene ID
<i>Gpnm</i>	Glycoprotein (transmembrane) nmb	8.88	4.6×10^{-9}	Mm.23567
<i>Sfrp1</i>	Secreted frizzled related sequence protein 1	5.74	1.4×10^{-7}	Mm.281691
<i>Gp49b</i>	Glycoprotein 49 B	4.82	4.4×10^{-6}	Mm.34408
<i>Ctss</i>	Cathepsin S	2.70	1.3×10^{-5}	Mm.3619
<i>Bex2</i>	Brain expressed X linked 2	0.26	1.3×10^{-5}	Mm.94160
<i>Col1a2</i>	Procollagen, type I, alpha 2	3.86	1.5×10^{-5}	Mm.277792
<i>Fcgr3</i>	Fc receptor, IgG, low affinity III	3.56	1.5×10^{-5}	Mm.22119
<i>Cd68</i>	CD68 antigen	3.46	1.5×10^{-5}	Mm.15819
<i>Fn1</i>	Fibronectin 1	3.51	1.8×10^{-5}	Mm.193099
<i>Itm2a</i>	Integral membrane protein 2A	2.10	1.8×10^{-5}	Mm.193
<i>Trem2b</i>	Triggering receptor expressed on myeloid cells 2b	4.00	2.0×10^{-5}	Mm.261623
<i>Ctsb</i>	Cathepsin B	2.93	2.6×10^{-5}	Mm.236553
<i>Bst1</i>	Bone marrow stromal cell antigen 1	2.20	2.6×10^{-5}	Mm.246332
<i>Col5a2</i>	Procollagen, type V, alpha 2	3.73	2.9×10^{-5}	Mm.10299
<i>Rnf149</i>	Ring finger protein 149	2.39	2.9×10^{-5}	Mm.28614

A similar analysis of microarray data was completed to determine the set of genes involved in the pulmonary response of a C3H mouse to bleomycin. The data from four arrays of treated mice (female 3 weeks, female 6 weeks; male 3 weeks, male 6 weeks) were compared to the data from two arrays generated from control mice (male control, female control). Substantially fewer genes ($n = 5$) in this strain were determined to be significantly over expressed in the lungs of treated mice compared to untreated controls and one gene was relatively under expressed (endothelial cell-specific molecule 1), which is consistent with the minimal lung response evident by histology. Of the genes increased in expression in the bleomycin treated C3H lung, glycoprotein (transmembrane) nmb, p21 (*cdkn1a*), and glycoprotein 49 B were also increased in expression in the B6 mouse lung after bleomycin, and to a significantly higher extent ($p < 0.05$) than in C3H (data not shown). Cathepsin K was the only gene identified to be induced by bleomycin in the lungs of C3H mice to a greater extent than in the lungs of B6 mice. The remaining two genes induced or repressed in the C3H lung (pleckstrin homology-like domain, family A, member 3, and endothelial cell-specific molecule 1), were similarly induced or repressed in the bleomycin treated B6 lungs (data not shown).

Strain comparison of gene expression data

Analyses were completed to identify the set of lung genes which differed in expression between B6 and C3H mice. The data from the treated mice (five B6 chips, four C3H chips) were compared and the data from untreated mice (two B6 control chips, two C3H control chips) were also evaluated. A total of 734 genes were determined to be significantly over expressed in the lungs of B6 bleomycin treated mice compared to C3H treated mice and 705 genes were relatively under expressed ($p < 0.05$; supplementary table 2, see <http://www.jmedgenet.com/supplemental>). In addition, 453 ESTs were differentially expressed ($p < 0.05$) in the lungs of bleomycin treated B6 mice compared to C3H (supplementary table 2).

The set of genes differentially expressed between bleomycin treated B6 and C3H mice was studied to identify those for which a strain difference in expression exists, independent of bleomycin treatment. The evaluation of the microarray data representing gene expression in the lungs of control mice revealed that 201 of the genes identified as differentially expressed in the lungs of bleomycin treated B6 and C3H mice were also differentially expressed in the control mice and for 176 (88%) of these the bleomycin treatment did not alter this strain difference in expression by

>2 fold (data not shown). Approximately half of the differentially expressed genes in this comparison of control datasets were found to be expressed at a higher level in B6 mice than in C3H mice.

For the majority of the genes differentially expressed between bleomycin treated B6 and C3H mice (89%), the differential expression was induced by the bleomycin treatment. By Gene Ontology analysis, the biological processes most often represented in the set of differentially expressed genes were, as in the B6 response to the drug, cell communication and response to stimulus. As shown in table 4, the most significantly differentially expressed genes included those involved in immune response such as the major histocompatibility complex gene H2-D1, complement component 4, and regenerating islet derived 3 γ ; and genes involved in protein transport including coatamer protein complex, subunit ζ 2, and solute carrier family 15 (H+/peptide transporter), member 2. A complete list of differentially expressed genes is available in supplementary table 2 (<http://www.jmedgenet.com/supplemental>). The datasets given in supplementary tables 1 and 2 indicate substantial gene expression profile differences between the B6 response to the drug relative to control (supplementary table 1) and the B6 relative to C3H bleomycin response (supplementary table 2). Approximately 53% of the genes of the B6 intrastain response dataset are not present in the interstrain response, and 56% of the B6/C3H differentially expressed genes are not of altered expression in the lungs of bleomycin treated B6 mice compared to controls.

Blmpf1 and 2 differential gene expression

To propose candidate genes for *Blmpf1* and 2, the set of genes which were both differentially expressed in the bleomycin treated lungs of B6 mice relative to C3H mice, and which map to the linkage intervals, was documented. Thirty seven differentially expressed genes mapped to *Blmpf1* and 19 to *Blmpf2*. The subset of this group, for which the gene expression of one strain was at least 1.5 fold of the other, is shown in table 5. The set of candidate genes includes major histocompatibility complex genes, oxidative stress response genes, and DNA damage response genes among others. In addition, six ESTs mapping to *Blmpf2* and five ESTs mapping to *Blmpf1* were identified to be differentially expressed in the lungs of bleomycin treated B6 mice relative to C3H mice, with expression changes ranging from 0.56 to 1.68 fold ($p < 0.05$; data not shown).

To further assess these candidate genes, the B6/C3H sequence comparison (described earlier) was reviewed. This analysis revealed 18 of the 56 candidate genes contained

Table 4 Genes most significantly differentially expressed in the lungs of bleomycin treated B6 mice relative to bleomycin treated C3H mice

Symbol	Name	Fold (B6/C3H)	p Value	UniGene ID
<i>H2-D1</i>	Histocompatibility 2, D region locus 1	0.06	2.8×10^{-11}	Mm.33263.1
<i>Rpgrip1</i>	Retinitis pigmentosa GTPase regulator interacting protein 1	6.50	5.4×10^{-10}	Mm.21662.1
<i>Lin7c</i>	Lin 7 homolog c (<i>C elegans</i>)	8.75	5.4×10^{-10}	Mm.235300
<i>Gbp1</i>	Guanylate nucleotide binding protein 1	0.14	1.0×10^{-9}	Mm.250.1
<i>Sfrp1</i>	Secreted frizzled related sequence protein 1	5.74	1.6×10^{-9}	Mm.281691
<i>C4</i>	Complement component 4 (within H-2S)	3.78	6.6×10^{-9}	Mm.16106.1
<i>Cap1</i>	CAP, adenylate cyclase associated protein 1 (yeast)	0.15	8.3×10^{-9}	Mm.8687.1
<i>Hba-a1</i>	Haemoglobin alpha, adult chain 1	0.17	9.9×10^{-9}	Mm.196110.1
<i>Reg3g</i>	Regenerating islet derived 3 gamma	6.23	5.9×10^{-8}	Mm.252385
<i>Prdx2</i>	Peroxiredoxin 2	2.89	6.0×10^{-8}	Mm.270130
<i>Copz2</i>	Coatamer protein complex, subunit zeta 2	2.25	1.1×10^{-7}	Mm.22144.1
<i>A1BG</i>	Alpha-1-B glycoprotein	0.36	2.0×10^{-7}	Mm.212712.2
<i>Paip1</i>	Polyadenylate binding protein interacting protein 1	0.34	2.0×10^{-7}	Mm.132584
<i>Pigo</i>	Phosphatidylinositol glycan, class O	2.16	2.2×10^{-7}	Mm.143738.2
<i>Slc15a2</i>	Solute carrier family 15 (H+/peptide transporter), member 2	2.27	2.2×10^{-7}	Mm.281804

coding region non-synonymous polymorphisms and/or 5' or 3' UTR SNPs, and 15 genes were without such changes (in at least 90% of the gene), while for 23 genes the C3H sequence was not available. The type of sequence variation identified for the differentially expressed candidate genes is listed in table 5.

To determine if the expression level of all the genes of the linkage regions had been evaluated, a review of the Affymetrix MOE430A chip was made using the Ensembl and Affymetrix databases. Approximately 55% of the genes in *Blmpf1* and 53% of the genes of *Blmpf2* are contained on this microarray. Of the 166 genes in *Blmpf2* and not on the Affymetrix chip, 103 (62%) are Riken cDNAs or Ensembl predicted genes, 35/166 (21%) are olfactory receptors, and 28 (17%) are known genes. Of the 150 genes of *Blmpf1* not found on the microarray used in these studies, 61 (41%) are Riken cDNAs or Ensembl predicted genes, 37 (25%) are olfactory receptors, and 52 (34%) are known genes (including 13 genes of the major histocompatibility complex).

To assess the validity of the gene expression differences identified by microarray analysis, the expression levels of four

genes were assessed with RT-PCR analysis. These data are presented in fig 3A. These genes were chosen to investigate examples of both differential and non-differential gene expression levels. As shown in fig 3B, the gene expression differences identified by microarray analysis were consistent with the results obtained through real time PCR.

DISCUSSION

A set of candidate genes was implicated in susceptibility to bleomycin induced pulmonary fibrosis through gene expression and sequence analysis combined with prior data from linkage studies. In prior studies, *Blmpf1* and 2 were mapped to 13 and 28 Mb intervals containing a total of 679 known or unknown genes. Using a combination of approaches, polymorphisms in specific genes of the oxidative stress response, apoptosis, and immune pathways have been identified as the potential causal variants in the development of fibrosis in this mouse model.

Consistent with previous studies of this model,^{7 11 16} the B6 mice developed substantial fibrosis in response to bleomycin exposure while C3H strain mice did not. The B6 response to

Table 5 B6:C3H differentially expressed genes of linkage regions *Blmpf1* and 2

Symbol	Name	Fold (B6/C3H)	P value	UniGene ID	Sequence variation*
<i>Blmpf1</i>					
<i>H2-D1</i> †	Histocompatibility 2, D region locus 1	0.06	2.8×10^{-11}	Mm.33263	Missense
<i>C4</i> †	Complement component 4 (within H-2S)	3.78	6.6×10^{-9}	Mm.16106	Missense
<i>Glo1</i> †	Glyoxalase 1	0.48	1.6×10^{-5}	Mm.261984	3' UTR
<i>H2-Q7</i> †	Histocompatibility 2, Q region locus 7	5.31	2.0×10^{-5}	Mm.296496	Unknown
<i>H2-K</i> †	Histocompatibility 2, K region	0.55	1.2×10^{-4}	Mm.244446	Missense‡
<i>Gabbr1</i> †	Gamma aminobutyric acid (GABA-B) receptor, 1	0.59	1.8×10^{-4}	Mm.245164	Unknown
<i>Tff2</i> †	Trefoil factor 2 (spasmolytic protein 1)	1.65	4.0×10^{-4}	Mm.1825	Unknown
<i>Notch4</i> †	Notch gene homolog 4 (<i>Drosophila</i>)	0.55	6.3×10^{-4}	Mm.173813	3' UTR†
<i>Fkbp5</i>	FK506 binding protein 5	2.91	2.9×10^{-3}	Mm.276405	Unknown
<i>Pim1</i>	Proviral integration site 1	1.74	0.01	Mm.269678	Unknown
<i>Psmb9</i>	Proteasome (prosome, macropain) subunit, beta type 9	0.62	0.01	Mm.16251	5' UTR†
<i>H2-Ea</i>	Histocompatibility 2, class II antigen E alpha	0.63	0.01	Mm.15680	None
<i>Hspa1a</i>	Heat shock protein 1A	1.54	0.01	Mm.275405	Missense
<i>Aif1</i>	Allograft inflammatory factor 1	1.56	0.01	Mm.10747	Unknown
<i>Cdkn1a</i>	Cyclin dependent kinase inhibitor 1A (P21)	1.64	0.02	Mm.195663	3' UTR
<i>Angptl4</i>	Angiopoietin-like 4	1.69	0.02	Mm.196189	3' UTR†
<i>Blmpf2</i>					
<i>Scgb3a1</i> †	Secretoglobulin, family 3A, member 1	2.28	3.1×10^{-6}	Mm.22802.1	None
<i>Trim1</i> †	Tripartite motif protein 16	2.43	3.8×10^{-5}	Mm.117087.2	3' UTR
<i>Pttg1</i> †	Pituitary tumour transforming 1	1.91	1.9×10^{-4}	Mm.6856.3	Unknown
<i>Gria1</i>	Glutamate receptor, ionotropic, AMPA1 (alpha 1)	0.55	6.5×10^{-4}	Mm.41612.1	Missense

*From the B6/C3H sequence comparison. None: no sequence variation detected in $\geq 90\%$ of sequence comparison; Unknown: no mismatches identified but $< 90\%$ sequenced compared; Missense: putative missense mutation identified; †differential expression detected between B6 and C3H mice, in both bleomycin treated and controls; ‡Celera Mouse SNP reference database (v3.6) where B6/C3H comparison unavailable. RT-PCR verification of microarray data.

the drug was histologically similar in both male and female mice sacrificed at three time points, as was the near null response of C3H mice at the times evaluated. For gene expression analysis, this permitted essentially five replicates of the B6 bleomycin response and four of the C3H response. Our study revealed the B6 pulmonary response to bleomycin included altered expression of extracellular matrix component and regulation genes, of inflammatory pathway components, and of DNA damage response genes, in agreement with previous studies of the B6 response^{15, 28} and with clinical data.²⁹ The genetic and mechanistic complexity of pulmonary fibrosis is evident in that nearly 1900 genes were identified to be differentially expressed in the B6 bleomycin treated lung compared to the C3H.

Combining the expression data with the linkage data revealed approximately 10% of genes/ESTs in the linkage regions were differentially expressed in the lungs of bleomycin treated B6 and C3H mice. The defined list of gene expression candidates is likely to be reasonably complete as the microarray used in these studies permitted the assessment of ~90% of the known genes in the linkage intervals, when olfactory receptor genes, which have no established pathway in pulmonary fibrosis, are not considered. The identified genes are considered candidates as the causal variation leading to the development of bleomycin induced pulmonary fibrosis with the assumption that the fibrosis causative gene is differentially expressed in the bleomycin treated lung. For approximately half of the linkage and expression candidate genes identified, listed in table 5, the difference in expression was measured in untreated mice as well. This may indicate that B6 mice have an inherent deficiency in enzyme levels of certain pathways that leaves

this strain susceptible to fibrosis, as opposed to an induced deficiency. This finding would be consistent with the rapid development of fibrosis in this strain, as the maximal fibrosis was evident 2 weeks after the completion of drug delivery.

The majority of the bleomycin induced pulmonary fibrosis candidate genes encode proteins involved in a DNA damage or oxidative stress response, in apoptosis, or in immune/inflammatory pathways. Each of these mechanisms has been implicated in the development of fibrosis. Bleomycin induced pulmonary fibrosis develops through a series of pathways initiated by the bleomycin binding to DNA, principally in alveolar epithelial cells, and inducing lesions through a free radical reactive complex.³⁰ The DNA damage and oxidative stress may lead to apoptosis of the epithelium³¹ and to a TH2 inflammatory response in the lung by increased major histocompatibility complex presentation of B7 and class II molecules on epithelial cells,³² activating T cells to a TH2, or fibrotic, phenotype.³³ The extent and duration of this inflammatory response may also be controlled, in part, through apoptosis.³⁴ Specific to the candidate genes proposed from this model, the B6 response to bleomycin may differ from the C3H response due to increased levels of bleomycin in the lungs of B6 mice. The expression of pituitary tumour transforming-1, which is a thiol protease inhibitor³⁵ that may act on the enzyme which degrades bleomycin, bleomycin hydrolase, was higher in B6 mice than C3H. Genes involved in immune responses (H2; secretoglobin, family 3A, member 1; complement component 4; allograft inflammatory factor 1) were also differentially expressed and thus the distinct fibrotic reactions of these strains could be due to differences in T cell regulation. The decreased levels of glyoxalase 1³⁶ and increased levels of p21 in the lungs of B6 mice may inhibit

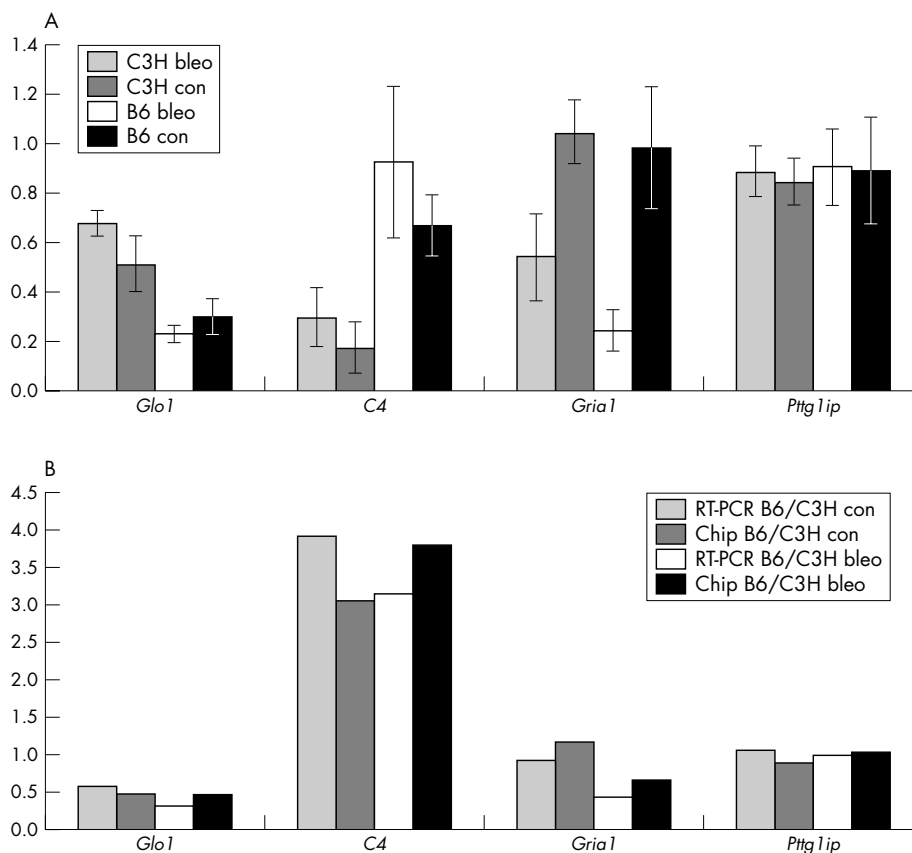


Figure 3 Comparison of microarray data with RT-PCR evaluation of selected genes. (A) The normalized expression of each target gene (\pm SE) for each strain and bleomycin treatment condition (bleo: treated; con: control). (B) The ratio of B6 to C3H gene expression level by microarray analysis and by RT-PCR analysis.

T cell apoptosis, and together with increased expression of proviral integration site-1, which induces T cell proliferation,³⁷ may have resulted in the maintenance of the profibrotic inflammatory signal. Finally, trefoil factor 2, which is increased in B6 mice, may promote the wound healing seen in fibrosis.³⁸

In addition to differential expression, the genes of the linkage regions were assessed for sequence variation. The NCBI database returned SNPs at a rate of 1/10 kbp, which is lower than that documented by others^{12–13} for the mouse genome, so the assessment of sequence variation was extended. Wade *et al*¹³ sequenced portions of the C3H genome in their study which identified the mouse genome as being comprised of regions of high and low SNP density. From the data of their report, both *Blmpf1* and 2 consist of random regions of low and high SNP density. Using these data and the publicly available B6 sequence, our analysis revealed up to 35% of the genes in the linkage regions contained coding sequence variations which would be predicted to alter the protein produced. The variation rate of 35% is likely an over estimate, due to the sequencing error rate of 1/1 kbp, and indeed an assessment of the Celera database, using B6 and A/J mice, revealed this type of SNP (coding non-synonymous) occurred in approximately 4% of the genes in *Blmpf1* or 2.

Sequence variation among inbred strains of known bleomycin induced fibrosis phenotype was used to assess candidate genes underlying *Blmpf1* and 2. Specifically, where C3H sequence is not yet known, SNPs between B6 and A/J (resistant) were reported. More generally, the evaluation of additional strains permits the reduction of background variation in the identification of causal variation. Only ~10% of the genes mapped to *Blmpf1* or 2 have non-intronic sequence variations predicting for the response of B6, A/J, and DBA/2 mice to bleomycin, including five genes of our candidate list. The inclusion of haplotypes of inbred strains other than B6 or C3H, however, is predicated on the assumption that the same loci, and the same genes, influence susceptibility to bleomycin induced pulmonary fibrosis in all evaluated strains. This is not known and no other loci for this trait have been mapped. The gene underlying *Blmpf1* may be relevant in this analysis as the MHC haplotype (within *Blmpf1*) has been shown to influence fibrosis susceptibility in other mouse strains,⁷ and this same region of chromosome 17 harbours loci of susceptibility to radiation induced pulmonary fibrosis³⁹ and ozone induced⁴⁰ and sulphate particle induced lung response.⁴¹

In previous studies, bleomycin hydrolase, the gene which encodes a protein that detoxifies bleomycin, was proposed as a candidate for *Blmpf2*.¹¹ The gene has subsequently been mapped to 76.5 Mb on chromosome 11, which is outside the linkage interval, and therefore was not evaluated for sequence differences. It was also not found to be differentially expressed in any comparison in this study. Also in the previous study, the influence of *Blmpf2* on the fibrosis phenotype of female mice was not clear as the linkage data identified this locus in male mice only, and the study of chromosome 11 consomic mice implicated *Blmpf2* in the response of mice of both sexes. In this study, there was no difference in gene expression between B6 male and female mice, thus it is presumed the *Blmpf2* locus alters the fibrosis response in mice of both sexes.

In summary we have used gene expression and DNA sequencing data to identify putative fibrosis susceptibility genes from a set of positional candidate genes. Through this approach, 10% of the positional candidate genes were measured to be differentially expressed in the bleomycin treated lung and DNA sequence analysis revealed approximately one third of the linked genes had B6/C3H sequence

variations which could affect the encoded protein. This subset of genes may be the most likely fibrosis candidate genes, although the positional candidate genes with sequence variations remain putative candidates. Additional testing to determine causality for a specific gene would include confirmation of sequence variation, target cell affected, and ultimately, an evaluation of the phenotype of mice with the B6 allele in C3H background mice, possibly involving a combination of knockout and transgenic replacement.⁴² This list of candidates should facilitate the identification of fibrosis susceptibility genes and the potential mechanistic interaction between the *Blmpf* loci.

ELECTRONIC-DATABASE INFORMATION



The supplementary tables are available at <http://www.jmedgenet.com/supplemental>. The following websites have also been mentioned in this article: Bioconductor: <http://www.bioconductor.org/>; TreeView: <http://rana.lbl.gov/>; Ensembl: http://www.ensembl.org/Mus_musculus/; NCBI: <http://www.ncbi.nlm.nih.gov/SNP/MouseSNP.cgi>; MIT: http://www.broad.mit.edu/ftp/pub/mouse_contigs/Mar10_02/; and Celera: <http://www.celera.com/>.

Authors' affiliations

C K Haston, T G Tomko, L Kerckhoff, Meakins-Christie Laboratories and Department of Medicine, McGill Centre for Bioinformatics, McGill University, Montreal, Canada

N Godin, M T Hallett, School of Computer Science, McGill University, Montreal, Canada

This work was supported by funding from the Canadian Institutes of Health Research and Fonds de la Recherche en Santé Québec.

Competing interests: none declared

Ethical approval: Animal use was approved by the McGill University Animal Care Committee and was in accordance with the guidelines of the Canadian Council on Animal Care.

REFERENCES

- Crystal RG**, Bitterman PB, Mossman B, Schwarz MI, Sheppard D, Almsay L, Chapman HA, Friedman SL, King TE Jr, Leinwand LA, Liotta L, Martin GR, Schwartz DA, Schultz GS, Wagner CR, Musson RA. Future research directions in idiopathic pulmonary fibrosis: summary of a national heart, lung, and blood institute working group. *Am J Respir Crit Care Med* 2002;**166**:236–46.
- Kolb M**, Bonnicaud P, Galt T, Sime PJ, Kelly MM, Margetts PJ, Gaudie J. Differences in the fibrogenic response after transfer of active transforming growth factor-beta 1 gene to lungs of "fibrosis-prone" and "fibrosis-resistant" mouse strains. *Am J Respir Cell Mol Biol* 2002;**27**:141–50.
- Gross TJ**, Hunninghake GW. Idiopathic pulmonary fibrosis. *N Engl J Med* 2001;**345**:517–25.
- Sime PJ**, O'Reilly KM. Fibrosis of the lung and other tissues: new concepts in pathogenesis and treatment. *Clin Immunol* 2001;**99**:308–19.
- Winkler MK**, Fowlkes JL. Metalloproteinase and growth factor interactions: do they play a role in pulmonary fibrosis? *Am J Physiol Lung Cell Mol Physiol* 2002;**283**:L1–11.
- Marshall RP**, McAnulty RJ, Laurent GJ. The pathogenesis of pulmonary fibrosis: is there a fibrosis gene? *Int J Biochem Cell Biol* 1997;**29**:107–20.
- Rossi GA**, Szapiel S, Ferrans VJ, Crystal RG. Susceptibility to experimental interstitial lung disease is modified by immune- and non-immune-related genes. *Am Rev Respir Dis* 1987;**135**:448–55.
- Schrier DJ**, Kunkel RG, Phan SH. The role of strain variation in murine bleomycin-induced pulmonary fibrosis. *Am Rev Respir Dis* 1983;**127**:63–6.
- Kawai K**, Akaza H. Bleomycin-induced pulmonary toxicity in chemotherapy for testicular cancer. *Expert Opin Drug Saf* 2003;**2**:587–96.
- Gabazza EC**, Taguchi O, Adachi Y. Bleomycin-induced lung fibrosis: the authors should have used another method to induce pulmonary lesions resembling human idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2002;**165**:845–6.
- Haston CK**, Wang M, Dejourmett RE, Zhou X, Ni D, Gu X, King TM, Weil MM, Newman RA, Amos CI, Travis EL. Bleomycin hydrolase and a genetic locus within the MHC affect risk for pulmonary fibrosis in mice. *Hum Mol Genet* 2002;**11**:1855–63.
- Waterston RH**, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Atwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F,

- Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyraas E, Felsenfeld A, Fewell GA, Flicke P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigo R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, Karolchik D, Kasprzyk A, Kawai J, Keibler E, Kells C, Kent WJ, Kirby A, Kolbe DL, Korf I, Kucherlapati RS, Kulbokas EJ, Kulp D, Landers T, Leger JP, Leonard S, Letunic I, Levine R, Li J, Li M, Lloyd C, Lucas S, Ma B, Maglott DR, Mardis ER, Matthews L, Mauceli E, Mayer JH, McCarthy M, McCombie WR, McLaren S, McLay K, McPherson JD, Meldrim J, Meredith B, Mesirov JP, Miller W, Miner TL, Mongin E, Montgomery KT, Morgan M, Mott R, Mullikin JC, Muzny DM, Nash WE, Nelson JO, Nhan MN, Nicol R, Ning Z, Nusbaum C, O'Connor MJ, Okazaki Y, Oliver K, Overton-Larty E, Pachter L, Parra G, Pepin KH, Peterson J, Pevzner P, Plumb R, Pohl CS, Poliakov A, Ponce TC, Ponting CP, Potter S, Quail M, Reymond A, Roe BA, Roskin KM, Rubin EM, Rust AG, Santos R, Sapojnikov V, Schultz B, Schultz J, Schwartz MS, Schwartz S, Scott C, Seaman S, Searle S, Sharpe T, Sheridan A, Shownkeen R, Sims S, Singer JB, Slater G, Smit A, Smith DR, Spencer B, Stabenau A, Stange-Thomann N, Sugnet C, Suyama M, Tesler G, Thompson J, Torrents D, Traversakis E, Tromp J, Ucla C, Ureta-Vidal A, Vinson JP, Von Niederhausern AC, Wade CM, Wall M, Weber RJ, Weiss RB, Wendl MC, West AP, Wetterstrand K, Wheeler R, Whelan S, Wierzbowski J, Willey D, Williams S, Wilson RK, Winter E, Worley KC, Wyman D, Yang S, Yang SP, Zdobnov EM, Zody MC, Lander ES; Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* 2002;**420**:520–62.
- 13 **Wade CM**, Kulbokas EJ 3rd, Kirby AW, Zody MC, Mullikin JC, Lander ES, Lindblad-Toh K, Daly MJ. The mosaic structure of variation in the laboratory mouse genome. *Nature* 2002;**420**:574–8.
- 14 **Chung MP**, Monick MM, Hamzeh NY, Butler NS, Powers LS, Hunninghake GW. Role of repeated lung injury and genetic background in bleomycin-induced fibrosis. *Am J Respir Cell Mol Biol* 2003;**29**:375–80.
- 15 **Kaminski N**, Allard JD, Pittet JF, Zuo F, Griffiths MJ, Morris D, Huang X, Sheppard D, Heller RA. Global analysis of gene expression in pulmonary fibrosis reveals distinct programs regulating lung inflammation and fibrosis. *Proc Natl Acad Sci U S A* 2000;**97**:1778–83.
- 16 **Haston CK**, Amos CI, King TM, Travis EL. Inheritance of susceptibility to bleomycin-induced pulmonary fibrosis in the mouse. *Cancer Res* 1996;**56**:2596–601.
- 17 **Perkowski S**, Sun J, Singhal S, Santiago J, Leikauf GD, Albelda SM. Gene expression profiling of the early pulmonary response to hyperoxia in mice. *Am J Respir Cell Mol Biol* 2003;**28**:682–96.
- 18 **Novak JP**, Sladek R, Hudson TJ. Characterization of variability in large-scale gene expression data: implications for study design. *Genomics* 2002;**79**:104–13.
- 19 **Ihaka R**, Gentleman R. R: a language for data analysis and graphics. *J Comput Graphical Stat* 1996;**5**:299–314.
- 20 **Irizarry RA**, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 2003;**31**:e15.
- 21 **Smyth GK**. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;**3**:article 3.
- 22 **Lönnstedt I**, Speed TP. Replicated microarray data. *Statistica Sin* 2002;**12**:31–46.
- 23 **Eisen MB**, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998;**95**:14863–68.
- 24 **Al-Shahrour F**, Díaz-Uriarte R, Dopazo J. FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics* 2004;**20**:578–80.
- 25 **Winer J**, Jung CK, Shackel I, Williams PM. Development and validation of real-time quantitative reverse transcriptase-polymerase chain reaction for monitoring gene expression in cardiac myocytes in vitro. *Anal Biochem* 1999;**270**:41–9.
- 26 **Lemon WJ**, Swinton CH, Wang M, Barbari N, Wang Y, You M. Single nucleotide polymorphism (SNP) analysis of mouse pulmonary adenoma susceptibility loci 1–4 for identification of candidate genes. *J Med Genet* 2003;**40**:e36.
- 27 **Yan Y**, Wang M, Lemon WJ, You M. Single nucleotide polymorphism (SNP) analysis of mouse quantitative trait loci for identification of candidate genes. *J Med Genet* 2004;**41**:e111.
- 28 **Katsuma S**, Nishi K, Tanigawara K, Ikawa H, Shiojima S, Takagaki K, Kaminishi Y, Suzuki Y, Hirasawa A, Ohgi T, Yano J, Murakami Y, Tsujimoto G. Molecular monitoring of bleomycin-induced pulmonary fibrosis by cDNA microarray-based gene expression profiling. *Biochem Biophys Res Commun* 2001;**288**:747–51.
- 29 **Zuo F**, Kaminski N, Eugui E, Allard J, Yakhini Z, Ben-Dor A, Lollini L, Morris D, Kim Y, DeLustro B, Sheppard D, Pardo A, Selman M, Heller RA. Gene expression analysis reveals matrilysin as a key regulator of pulmonary fibrosis in mice and humans. *Proc Natl Acad Sci U S A* 2002;**99**:6292–7.
- 30 **Ramotar D**, Wang H. Protective mechanisms against the antitumor agent bleomycin: lessons from *Saccharomyces cerevisiae*. *Curr Genet* 2003;**43**:213–24.
- 31 **Inoshima I**, Kuwano K, Hamada N, Yoshimi M, Maeyama T, Hagimoto N, Nakanishi Y, Hara N. Induction of CDK inhibitor p21 gene as a new therapeutic strategy against pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol* 2004;**286**:L727–33.
- 32 **Kaneko Y**, Kuwano K, Kunitake R, Kawasaki M, Hagimoto N, Hara N. B7-1, B7-2 and class II MHC molecules in idiopathic pulmonary fibrosis and bronchiolitis obliterans-organizing pneumonia. *Eur Respir J* 2000;**15**:49–55.
- 33 **Mastruzzo C**, Crimi N, Vancheri C. Role of oxidative stress in pulmonary fibrosis. *Monaldi Arch Chest Dis* 2002;**57**:173–6.
- 34 **Hao Z**, Hampel B, Yagita H, Rajewsky K. T cell-specific ablation of Fas leads to Fas ligand-mediated lymphocyte depletion and inflammatory pulmonary fibrosis. *J Exp Med* 2004;**199**:1355–65.
- 35 **Bernal JA**, Luna R, Espina A, Lazaro I, Ramos-Morales F, Romero F, Arias C, Silva A, Tortolero M, Pintor-Toro JA. Human securin interacts with p53 and modulates p53-mediated transcriptional activity and apoptosis. *Nat Genet* 2002;**32**:306–11.
- 36 **Sakamoto H**, Mashima T, Kizaki A, Dan S, Hashimoto Y, Naito M, Tsuruo T. Glyoxalase I is involved in resistance of human leukemia cells to antitumor agent-induced apoptosis. *Blood* 2000;**95**:3214–8.
- 37 **Mikkers H**, Nawijn M, Allen J, Brouwers C, Verhoeven E, Jonkers J, Berns A. Mice deficient for all PIM kinases display reduced body size and impaired responses to hematopoietic growth factors. *Cell Biol* 2004;**24**:6104–15.
- 38 **Nikolaidis NM**, Zimmermann N, King NE, Mishra A, Pope SM, Finkelman FD, Rothenberg ME. Trefoil factor-2 is an allergen-induced gene regulated by Th2 cytokines and STAT6 in the lung. *Am J Respir Cell Mol Biol* 2003;**29**:458–64.
- 39 **Haston CK**, Travis EL. Murine genetic susceptibility to radiation-induced pulmonary fibrosis is influenced by a genetic factor implicated in susceptibility to bleomycin-induced pulmonary fibrosis. *Cancer Res* 1997;**57**:5286–91.
- 40 **Kleeberger SR**, Levitt RC, Zhang LY, Longphre M, Harkema J, Jedlicka A, Eleff SM, DiSilvestre D, Holroyd KJ. Linkage analysis of susceptibility to ozone-induced lung inflammation in inbred mice. *Nat Genet* 1997;**17**:475–8.
- 41 **Kleeberger SR**. Genetic aspects of susceptibility to air pollution. *Eur Respir J Suppl* 2003;**40**:52s–56s.
- 42 **Glazier AM**, Nadeau JH, Aitman TJ. Finding genes that underlie complex traits. *Science* 2002;**298**:2345–9.