Genomic and Genome-wide Association of Susceptibility to Radiation-induced Fibrotic Lung Disease in Mice

Running Head: Association to Radiotherapy Fibrosis

Alexandra Paun[1] and Christina K. Haston[1,2]

[1]Meakins-Christie Laboratories and the Departments of Human Genetics and [2]Medicine, McGill University, Montreal, PQ, Canada

Corresponding Author*:* Christina K. Haston, Ph.D., Meakins-Christie Laboratories, Department of Medicine, McGill University, 3626 St. Urbain, H2X 2P2, Montreal, PQ, Canada. e-mail: christina.haston@mcgill.ca; Tel: (514) 398-3864 x.089714. Fax: (514) 398-7483.

total number of pages = 26

total number of tables and figures = 5

Keywords: Quantitative Trait Loci, mouse SNP map; pneumonitis; fibrosis; radiation response

Abstract

Background and Purpose: To identify genes which influence the fibrotic response to thoracic cavity radiotherapy, we combined a genome wide single nucleotide polymorphism (SNP) association evaluation of inbred strain response with prior linkage and gene expression data.

Material and Methods: Mice were exposed to 18 Gy whole thorax irradiation and survival, bronchoalveolar cell differential, and histological alveolitis and fibrosis phenotypes were determined. Association analyses were completed with 1.8 million SNPs in single markers and haplotypes.

Results: Nine strains developed significant fibrosis and 11 strains succumbed to alveolitis only or alveolitis with minimal fibrosis. Post irradiation survival time ($p<0.001$) and bronchoalveolar lavage neutrophil percent ($p=0.055$) were correlated with extent of alveolitis and were not significantly correlated with fibrosis. Genome wide SNP analysis identified 10 loci as significantly associated with radiation-induced fibrotic lung disease ($p<8.41 \times 10^{-6}$; by permutation test), with the most significant SNP within a conserved non-coding region downstream of cell adhesion molecule 1 (*Cadm1)*. Haplotype and SNP analyses performed within previously-identified loci revealed additional genes containing SNPs associated with fibrosis including *Slamf6* and *Cdkn1a*.

Conclusions: Combining genomic approaches identified variation within specific genes which function in the tissue response to injury as associated with fibrosis following thoracic irradiation in mice.

Introduction: The common cancer treatment modality of radiotherapy, when delivered to the thoracic cavity can produce serious inflammatory (alveolitis or pneumonitis) or fibrotic side effects in the lung [1-4]. Patients who are likely to develop fibrosis, an incurable normal tissue response to radiation, at present can not be identified before therapy [1]. Efforts to isolate specific genetic variants which are associated with clinical radiation response have focussed on candidate gene studies of small sample sized populations and have not produced consistent results [5]. In addition, the approach of using clinical data alone to identify causal genetic variations influencing susceptibility to radiation-induced lung disease can be confounded by the effects of multiple genes and their interactions on the phenotype, variability in the clinical description of the phenotype and by the cancer and other therapies.

To circumvent these limitations, inbred mouse strains which vary in their propensity to develop alveolitis and fibrosis after radiation exposure [6] can be evaluated to isolate candidate genetic variation contributing to the lung response phenotype. Indeed, the clinical outcomes of thoracic radiotherapy, with respect to adverse effects, of fibrosis and alveolitis, and the times at which they occur, are each represented in the radiation response of an inbred mouse strain. We [7-10] and others [6, 11, 12] have reported murine strain differences in response to radiation wherein following high dose whole thorax irradiation, C3H/HeJ (C3H) mice develop a diffuse lethal alveolitis, characterized by an inflammatory infiltrate of alveoli, and C57BL/6J (B6) mice respond to lung irradiation with fibrosis. The phenotypic difference between B6 and C3H mice has been used in replicate studies [7, 13] to map loci of radiation-induced pulmonary fibrosis, named *Radpf1* on chromosome 17 and *Radpf2* on chromosome 1. Additional loci on chromosome 6 [ref 13] and chromosomes 7, 9 [ref 7] were detected in either one of the studies only, while the strain difference in alveolitis has been mapped to loci distinct from those of

fibrosis [7].

The similarity to the clinical phenotype presented by the mouse models enables use of mouse genomic resources to investigate the genetic basis of radiation-induced lung disease. Among these is the mouse single nucleotide polymorphism or SNP map [14, 15] which has been used to complete genome-wide association studies (GWAS) in inbred mice for traits such as acetaminophen toxicity [16], acute lung disease [17] and lung cancer [18-20]. Mapping of disease traits in mice through association data alone however can produce spurious results due to the relatedness of inbred strains created by their unique breeding history [21] and this approach may have reduced statistical power to detect associations given the limited number of inbred strains. To compensate, specific analyses incorporating inbred strain relatedness have been developed [21] and association data are often combined with quantitative trait loci determined through mapping or through the evaluation of specific congenic or consomic mice for the purpose of identifying causal genetic variation associated with the trait [22]. Association analysis limited to QTL regions has also been used to identify candidate quantitative trait genes [23, 24]. The second mouse genetics tool of chromosome substitution, or consomic mice, which are mice bred to contain one specific donor chromosome in a recipient inbred strain background [25], enables the assessment of a genomic interval contribution to a trait without the production of congenic mice.

In this study, we combine genomics approaches to identify specific genetic variants associated with predisposition to radiation-induced pulmonary fibrosis. Initially, to determine whether loci of lung response mapped in B6xC3H F2 [13] and backcross [7] mice might predict for response in another strain we made use of the difference in radiation response between B6 (fibrosis) and A/J (alveolitis) mice [8] and we evaluated the radiation response of available

B6:A/J consomic mice. We next defined the radiation response of a panel of inbred strains and these data, combined with that from a high density SNP dataset were used to identify genes with allelic variants shared among alveolitis (only) responding mice which were distinct from those shared by fibrosis responding mice. Lastly we integrated data from a gene expression profiling experiment [9] and from an assessment of species sequence conservation to refine the list of associated genes.

Materials and Methods

(see Supplementary Data for more detailed Materials and methods)

Mice: Mice of inbred and consomic B6.A strains were purchased from the Jackson Laboratory (Bar Harbor, ME, USA) and housed in the animal facility of the Meakins-Christie Laboratories. All mice were handled according to guidelines and regulations of the Canadian Council on Animal Care.

Radiation treatment: Lung damage was elicited by whole thorax radiation exposure (18 Gray; dose rate 0.7 Gray/minute) using a Gamma cell Cesium-137 unit as previously described [7-10]. The irradiated mice were sacrificed when moribund, or at 26 weeks after treatment. The control mice were not treated and were euthanized at the 15-26 week time points.

Histology: At necropsy, bronchoalveolar lavage (BAL) was performed by cannulating the trachea and retrieving cells from one 1-mL injection of phosphate buffered saline. The lungs were then removed and the left lobe of each mouse perfused with 10% neutral buffered formalin and submitted for histological processing. Lung sections were stained with Masson's Trichrome

and the area of fibrosis in the left lung lobe was determined from a user drawn region surrounding the fibrosis (Image Pro Plus Software) compared to the area of the entire lobe to yield the percent of pulmonary fibrosis for individual mice [7,8]. To assess alveolitis, hematoxylin and eosin stained left lung sections were evaluated through semi-quantitative histology [7, 8]. Mast cell numbers were determined by counting the positive cells present in 10 fields of a lung histological section stained with Toluidine blue, at a 400x magnification [10]. All scoring was completed by a user blinded to mouse strain and treatment.

Bronchoalveolar Lavage Fluid (BAL) Analysis: Inflammatory cell counts were performed (400x magnification) on cytocentrifuged cells (214.2g for three minutes), after staining with a hematoxylin-eosin kit and are reported as the percentage of 500 counted cells. To assess correlations among disease phenotypes pairwise Pearson's correlation coefficients and corresponding p-values were calculated using the rcor.test function in R (http://cran.r-project.org/).

Genome-wide association studies (GWAS): GWAS were completed using 236 mice from 27 inbred strains and the following phenotypes: percent fibrotic lung tissue, survival post irradiation, percent of lymphocytes & neutrophils in BAL and lung tissue mast cell count. The phenotype data include the response of one SWR/J mouse, which was included as the measured phenotype agrees with that previously reported for this strain [6].

Genotype data consisted of the CGD1 (http://cgd.jax.org/index.php) imputed dataset downloaded from the Mouse Phenome Database (Jackson Laboratory, Bar Harbor, ME). This dataset was filtered for SNPs imputed with confidence scores >0.9 as in [18], and SNPs with

fewer than 20 strains typed/imputed or with a minor allele frequency <3/27 were removed as has been used by others [18,19]. The 1,885,108 remaining SNPs span the mouse genome at an average of 1 per 1.8 kB and were assessed for association through single marker analysis and haplotype analysis.

In single marker analysis genotype-phenotype associations were evaluated for each individual SNP using efficient mixed model association [EMMA, ref 21], as used by others [18, 26]. The significance threshold was determined through permutation testing as in [19,26].

QTL-wide association studies (QWAS): To specifically assess the SNP variation within the previously identified linkage intervals [7, 13] for association with radiation-induced lung disease we performed QWAS. For this SNPs located in linkage regions on chromosomes 1,6,7,9 & 17 [refs 7,13] and the percent fibrosis phenotype data of the GWAS were re-analysed with the single marker method. We added the condition that the A/J and C3H alleles be different from the B6 allele for the regions on chromosomes 1 & 17; and that the C3H allele be different from the B6 allele for the regions on chromosomes 6, 7 & 9. The significance level for association was determined through permutation testing with the QTL region SNPs and the 5^{th} percentile of the resultant distribution was taken as the QTL-wide significance threshold.

Linkage interval genotypes were also tested for association with the fibrosis phenotype using haplotype blocks derived from the SNP data through a build-and-step-back approach. Our algorithm iteratively joins consecutive SNPs into haplotype blocks by clustering strains based on their genotypes. In this analysis each haplotype block was represented by a string of "labels" for each strain indicating the haplotype belonging to each strain. The association of the phenotype (percent fibrosis) with the haplotype "labels" was assessed using an analysis of variance

(ANOVA). The significance threshold was determined through a weighted bootstrap procedure which takes into account strain relatedness while correcting for multiple testing. However this method was too stringent ($<10^{-32}$) and yielded no significant results. We therefore present the ten most significant haplotypes in each QTL as in [27], for which the lowest nominal p-values ranged from 10^{-12} for chr 7 to 10^{-19} for chr 9.

SNP-SNP interaction between Radpf1 and Radpf2: To determine whether an interaction of genotypes within the linkage intervals *Radpf1* on chromosome 17 and *Radpf2* on chromosome 1 contributes to the fibrosis phenotype, as suggested in [13], we tested the association between the phenotype and all possible pairwise sets of SNPs as predictors through a linear model including a SNP-SNP interaction term. The significance of the interaction term was determined using an F test by comparing a full model including the interaction term with the partial model that excludes the interaction variable. Approximately 8 million tests were performed and the results were corrected for multiple testing using the Benjamini-Hochberg FDR approach.

Candidate Gene Identification: Genes with SNP variation meeting the GWAS/QWAS significance criteria were defined using BioMart feature of Ensembl, filtered for pulmonary expression, and the predicted SNP effect on the encoded protein or on gene expression was evaluated.

Results

B6.A consomic strain response to thoracic irradiation: To determine whether loci linked to the radiation-induced pulmonary fibrosis response in B6 mice relative to C3H mice also predict for

the radiation response of the B6 relative to A/J strain studies of B6.A consomic mice were completed. As shown in Figure 1A, consomic B6 mice with A/J alleles for chromosomes 1 or 17 developed minimal fibrosis, which was not significantly different from that of the A/J strain ($p>0.11$) following radiation treatment, while mice consomic for chromosomes 4, 14, or X, selected to be distinct from the *Radpf* loci, were not spared fibrosis, thus the previously mapped *Radpf* loci are also relevant to the B6:A/J strain comparison for this phenotype. Further phenotyping showed the mice of the B6.17A consomic strain to be resistant to the development of radiation-induced alveolitis, compared to the parental strains (Figure 1B,C).

Inbred strain response to thoracic irradiation: To enable the use of a mouse SNP map for evaluating genetic susceptibility to radiation-induced fibrosis, the radiation response to 18 Gy whole thorax irradiation of a panel of 27 inbred strains was determined. The majority of mice suffered from alveolitis in response to thoracic irradiation, evident through increased average histological scores compared to those of untreated mice, as shown in Figure 2. Only male mice of the BTBRt+ tf/J strain did not exhibit respiratory distress and thus were euthanized at the experimental endpoint of 26 weeks. Nine mouse strains also developed fibrosis wherein fibrotic scars covered greater than 4.5% of the lung (Figure 2).

The inflammatory response to irradiation was assessed by quantifying mast cell counts in lung tissue and bronchoalveolar cell differentials in mice euthanized in distress or at the end of the experiment. Neither measure was significantly correlated with the development of fibrosis in the inbred strain panel (Figure 3 and Supplementary Figure 1).

GWAS analysis of inbred strain radiation response: To determine whether specific polymorphisms are significantly associated with the lung response phenotypes single marker association analyses were completed with a genome wide SNP map. Regions on ten chromosomes were identified as associated with the development of the radiation-induced fibrosis lung response in the 27 inbred strains, $p<0.05$, as shown in Figure 4A. Significant associations were also detected for time to onset of respiratory distress (survival), mast cell count and lavage lymphocyte percent & neutrophil percent (Supplementary Figure 2).

QWAS analysis of inbred strain radiation response: To investigate variation within previously identified linkage intervals as associated with fibrosis susceptibility we performed an association analysis limited to five QTL intervals [7,13]. Using the single marker method the most significant associations were found for SNPs in the chromosome 9 locus (Figure 4B), followed by those on chromosomes 6 and 7.

Within the chromosome 1 (*Radpf2*) & 17 (*Radpf1*) linkage intervals no SNPs reached the QTL-wide threshold for significance. Given the previously suggested genetic interaction between variation within *Radpf1* & *Radpf2* on fibrosis susceptibility [13], we tested for a significant interaction of SNP genotypes at these loci and the phenotype of fibrosis. Assessing the interaction term between all pairs of SNPs in the two loci produced 10 significant SNP-SNP interactions (Supplementary Table 1).

To enable an evaluation of strain haplotypes within the linkage intervals as associated with fibrosis we developed an algorithm which iteratively "steps-back" to produce large, overlapping blocks of shared haplotypes among strains with a minimal error rate. The significantly fewer haplotype blocks in each QTL, compared to the numbers of SNPs, resulted in

a 5-10 fold reduction in the number of association tests performed for each QTL. An analysis of variance was used to reveal linkage interval haplotype blocks most significantly associated with radiation-induced pulmonary fibrosis. Since none of the blocks passed the stringent weighted bootstrap significance threshold, the 10 unique most significant p-values in each QTL were retained for candidate gene analysis. The most significant unique p-values correspond to an average of 21 blocks per QTL, and these overlap with a total of 8 genes, including 3 common to the single SNP QWAS set.

Candidate genes for susceptibility to radiation-induced pulmonary fibrosis: A summary of the candidate genes revealed by each analysis, filtered for expression in lung tissue, together with information of the nature of the polymorphisms and of the differential expression status of the genes, is given in Table 1. Based on the genome wide analyses ten candidate genes for radiation-induced pulmonary fibrosis were revealed while from the QTL level eight additional genes were identified as having variation which could influence the trait. Supplementary Figure 3 shows the effect of genotype on the pulmonary fibrosis phenotype for four of these significant associations presented in Table 1.

In addition, the *Radpf1-Radpf2* interaction evaluation revealed potential epistasis between two genes on chromosome 1 and six genes on chromosome 17 as affecting fibrosis and these are included as candidates in Table 1. As an example of the detected interaction, inbred strains with an allele "A" at SNP rs31529220, within *Slamf6* on chromosome 1, developed an average of 4% fibrosis in their lungs post irradiation, but, when the responses of these strains are segregated by genotypes at SNP rs31528128, near *Btnl1* on chromosome 17, average fibrosis scores of approximately 2% and 7%, depending on *Btnl1* SNP genotype, are revealed, as illustrated in

Supplementary Figure 4.

Discussion

Radiation-induced pulmonary side effects remain a significant clinical concern [2-4], thus research to reveal assays predictive of response to radiotherapy remains at the forefront of molecular radiation oncology [28-30]. Recent evidence indicates that interpatient variation in the development of side effects after cancer radiotherapy can not be explained, to a clinically relevant effect, by polymorphisms in individual candidate genes [31], therefore alternative approaches to investigating the genetic basis of radiotherapy side effects are needed [32]. One approach involves evaluating the whole genome, which allows the discovery of a priori unknown variants as associated with the trait, in the controlled environment of a model system.

In this study we made use of a mouse model which presents a phenotype similar (in time of onset and histologically) to the clinical radiotherapy late effect of pulmonary fibrosis and report the first genome-wide evaluation of radiation-induced side effects; a prevalent clinical problem [2-4]. Through association and linkage-based studies novel genetic variants which influence susceptibility to radiation-induced lung disease, including in genes which function in immune pathways, were revealed.

Using chromosome substitution (consomic) mice we have shown the *Radpf* loci, previously mapped in a B6 x C3H mouse cross [7, 13], to be relevant to the B6 strain response to radiation, relative to A/J, as A/J alleles on chromosomes 1 or 17 protected against fibrosis development. In addition, for chromosome 17, the presence of A/J alleles reduced the extent of radiation-induced alveolitis in consomic mice compared to B6 mice. This result is similar to that reported for the radiation response of congenic B6.AKR-H2^{k} mice [13] where the presence of AKR alleles in a genomic region of chromosome 17 enabled the post treatment survival of the

congenic B6 mice to the end of the experiment with minimal lung disease. These studies indicate that B6 alleles on chromosome 17 are necessary for the development of fibrosis, in B6 mice, and also contribute to the alveolitis response as the consomic B6.17A and congenic B6.AKR-H2^{k} mice did not develop the alveolitis presented by the inbred B6 mice.

From the 27 strains of mice we observed a clear phenotype separation in propensity to develop fibrosis following thoracic radiotherapy which supports the use of inbred mice as a tool to dissect this complex clinical trait. Our results showed the lung disease phenotype of B6, C3H, A/J, BALB/c, KK/HIJ, AKR/J, SWR and CBA mice to agree with responses presented in prior reports on these mice [6, 10, 11, 33], while to our knowledge, the response of the 20 remaining strains to pulmonary irradiation has not been previously reported. In addition to the fibrosis phenotype, strain dependent inflammatory responses to thoracic irradiation have also been documented [12, 34] including mast cell numbers and bronchoalveolar lavage cell differential [7]. The survey of radiation response in inbred mice indicates that the inflammatory phenotypes evaluated, bronchoalveolar lavage cell profile and mast cell counts, did not, however, correlate with the development of fibrosis, suggesting the inflammatory contributions to the trait, if any, are not the same in every inbred strain. Finally, consistent with prior datasets [7, 8] is the observation that time to respiratory distress was not predictive of the phenotype of fibrosis in the inbred strain panel.

Using the phenotype data of the inbred strain panel we performed GWA studies and identified novel genetic variants associated with susceptibility to radiation-induced pulmonary fibrosis. Our results are derived from the phenotypes of 27 inbred strains, a number which is reasonable for such studies [18, 20, 35] but includes limitations. For example, as with all association studies, a larger sample size would increase the statistical power to detect

polymorphisms with small effects. Secondly, our source for sequence polymorphism was the CGD1 dataset which contains both experimentally-derived SNPs and imputed SNPs and is, to our knowledge, the most dense genotype dataset available, pending the delivery of complete genomic sequence, which has been reported for a subset of our evaluated inbred strains [36]. Finally, the phenotypes and genotypes were analysed with EMMA, an algorithm which addresses strain relatedness, but does not permit genotypes to be evaluated in haplotypes.

We extended the linkage region association analysis, or QWAS, to include an assessment of haplotypes for association with the trait, given the limited genetic information of single marker analysis which can not reveal linked, interacting SNPs within disease loci. To facilitate this analysis haplotypes of a specified minimal error rate were generated. Tests for associations of fibrosis with extended haplotypes, however, produced no results exceeding the stringent threshold for significance, as we were unable to adequately account for strain relatedness in this analysis. Nevertheless by focusing on most significant associations, we identified the haplotype of several QTL genes which were found, through the single marker analysis, to be significantly associated with fibrosis and identified additional candidates for the trait.

Our investigation identified 26 genes as potential candidates for the lung response phenotype. The majority of the genes have been demonstrated to be differentially expressed in lung tissue following whole thorax irradiation [9], but, as is the case with most GWAS, many of significant SNPs found by our analysis reside in introns which makes the mechanism of their genetic contribution to disease difficult to interpret [37]. For example, on chr 9, highly significant SNPs and haplotypes were found in a conserved noncoding sequence of an intergenic region, between *Cadm1* and *Rexo2* closest to a cluster of *Fam55* genes and predicted genes/pseudogenes. Whether such variation could act in gene regulation, which is possible for

long-range cis-acting SNPs of gene targets [19, 38], and if so which of the nearby genes it acts upon however are uncertain. *Cadm1* is cell adhesion molecule 1 and as our previous study [9] demonstrated cell adhesion molecule signalling to be the pathway most altered in the lungs of mice following thoracic irradiation, variation in cell adhesion mediated by *Cadm1* could alter the lung response to radiation injury. Indeed, Hallahan *et al.* [39] showed intercellular adhesion molecule-1 deficient mice to develop less pulmonary fibrosis following thoracic irradiation than that measured in wild type controls. Secondly, *Rexo2* is an RNA exonuclease and recently another similar enzyme, *Xrn2* has been associated with murine lung cancer [19] through its involvement in epithelial cell proliferation. Impaired restoration of the alveolar epithelium following the initial stimulus is a potential trigger for pulmonary fibrosis [40, 41], therefore variation in a gene implicated in alveolar cell proliferation could contribute to radiation-induced lung damage.

Additional candidate genes revealed here by genomic investigation are also physiologically plausible for affecting the phenotype. For example, pulmonary fibrosis can occur as a result of increased extracellular matrix deposition and genes involved in this process such as *Abl2* and *Ddr1* [42,43] were identified to have variation significantly associated with the trait. Supporting this, *Ddr1* -/- mice have been shown to develop a reduced amount of pulmonary fibrosis in response to a related challenge, that of bleomycin-induced lung injury [44], and in separate work [45], to be protected from renal fibrosis. An inflammatory contribution to the tissue response of fibrosis has also been suggested [46] and polymorphisms within particular candidate genes which function to alter the inflammatory response (*Slamf6*, *Btnl1*) could thus affect susceptibility to radiation-induced lung disease [41, 47, 48]. Further, increased levels of the cytokine transforming growth factor beta (Tgf-β) in lung tissue have been implicated in the

development of fibrosis in animal models [49], including in response to irradiation [50] and specific candidate genes identified in our analysis (*Cdkn1a* or *p21, Pin1*) have been demonstrated to influence fibrosis development by altering Tgf-β levels [51, 52]. Finally, among the identified candidate genes *Shprh* is known to function in DNA repair [53], therefore variation in cellular radiation sensitivity such as reported by others [27, 54], in addition to strain specific tissue responses to radiation injury, may contribute to the development of the fibrosis phenotype.

The radiotherapy-associated side effect of pulmonary fibrosis is a genetically complex trait, whose development is most likely dictated by several genes which each contribute a small effect to the disease phenotype. Our investigation of the genetic basis of this trait therefore encompassed genome-wide and QTL-specific association analyses, both at the single marker and haplotype level, to increase the power of identifying true candidates for this trait. These data, combined with strain dependent and post thoracic irradiation gene expression profiling, have revealed variants in genes including cell adhesion molecules and those which function in adaptive immune responses to be associated with radiation-induced pulmonary fibrosis in mice.

Acknowledgements: These studies were funded by the Canadian Institutes of Health Research (#MOP62846) and Fonds de la Recherche en Sante Quebec, to CKH. The study sponsors did not contribute to the study design, the collection, analysis and interpretation of data; the writing of the manuscript; nor in the decision to submit the manuscript for publication.

Conflict of Interest Statement: None to disclose.

Reference List

[1] Graves PR, Siddiqui F, Anscher MS Movsas B. Radiation pulmonary toxicity: from mechanisms to management. Semin Radiat Oncol 2010;20:201-207.

[2] Kristensen CA, Nøttrup TJ, Berthelsen AK, et al. Pulmonary toxicity following IMRT after extrapleural pneumonectomy for malignant pleural mesothelioma. Radiother Oncol 2009;92:96-9.

[3] Guckenberger M, Baier K, Polat B, et al. Dose-response relationship for radiation-induced pneumonitis after pulmonary stereotactic body radiotherapy. Radiother Oncol 2010;97:65-70.

[4] Borst GR, Ishikawa M, Nijkamp J, et al. Radiation pneumonitis in patients treated for malignant pulmonary lesions with hypofractionated radiation therapy. Radiother Oncol 2009;91:307-13.

[5] Barnett GC, West CM, Dunning AM, et al. Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype. Nat Rev Cancer 2009;9:134-142.

[6] Sharplin J Franko AJ. A quantitative histological study of strain-dependent differences in the effects of irradiation on mouse lung during the intermediate and late phases. Radiat Res 1989;119:15-31.

[7] Haston CK, Begin M, Dorion G Cory SM. Distinct loci influence radiation-induced alveolitis from fibrosing alveolitis in the mouse. Cancer Res 2007;67:10796-10803.

[8] Lemay AM Haston CK. Radiation-induced lung response of AcB/BcA recombinant congenic mice. Radiat Res 2008;170:299-306.

[9] Paun A, Lemay AM Haston CK. Gene expression profiling distinguishes radiation-induced fibrosing alveolitis from alveolitis in mice. Radiat Res 2010;173:512-521.

[10] Thomas DM, Fox J Haston CK. Imatinib therapy reduces radiation-induced pulmonary mast cell influx and delays lung disease in the mouse. Int J Radiat Biol 2010;86:436-444.

[11] Jackson IL, Vujaskovic Z Down JD. Revisiting strain-related differences in radiation sensitivity of the mouse lung: recognizing and avoiding the confounding effects of pleural effusions. Radiat Res 2010;173:10-20.

[12] Chiang CS, Liu WC, Jung SM, et al. Compartmental responses after thoracic irradiation of mice: strain differences. Int J Radiat Oncol Biol Phys 2005;62:862-871.

[13] Haston CK, Zhou X, Gumbiner-Russo L, et al. Universal and radiation-specific loci influence murine susceptibility to radiation-induced pulmonary fibrosis. Cancer Res 2002;62:3782-3788.

[14] Frazer KA, Eskin E, Kang HM, et al. A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. Nature 2007;448:1050-1053.

[15] Szatkiewicz JP, Beane GL, Ding Y, Hutchins L, Pardo-Manuel de Villena F Churchill GA. An imputed genotype resource for the laboratory mouse. Mamm Genome 2008;19:199-208.

[16] Harrill AH, Watkins PB, Su S, et al. Mouse population-guided resequencing reveals that variants in CD44 contribute to acetaminophen-induced liver injury in humans. Genome Res 2009;19:1507-1515.

[17] Leikauf GD, Concel VJ, Liu P, et al. Haplotype association mapping of acute lung injury in mice implicates activin a receptor, type 1. Am J Respir Crit Care Med 2011;183:1499-1509.

[18] Liu PY, Vikis H, James M, et al. Identification of Las2, a major modifier gene affecting the Pas1 mouse lung tumor susceptibility locus. Cancer Res 2009;69:6290-6298.

[19] Lu Y, Liu P, James M, et al. Genetic variants cis-regulating Xrn2 expression contribute to the risk of spontaneous lung tumor. Oncogene 2010;29:1041-1049.

[20] Berndt A, Cario CL, Silva KA, et al. Identification of fat4 and tsc22d1 as novel candidate genes for spontaneous pulmonary adenomas. Cancer Res 2011;71:5779-5791.

[21] Kang HM, Zaitlen NA, Wade CM, et al. Efficient control of population structure in model organism association mapping. Genetics 2008;178:1709-1723.

[22] Manenti G, Galvan A, Pettinicchio A, et al. Mouse genome-wide association mapping needs linkage analysis to avoid false-positive Loci. PLoS Genet 2009;5:e1000331.

[23] Hillebrandt S, Wasmuth HE, Weiskirchen R, et al. Complement factor 5 is a quantitative trait gene that modifies liver fibrogenesis in mice and humans. Nat Genet 2005;37:835-843.

[24] Nissenbaum J, Devor M, Seltzer Z, et al. Susceptibility to chronic pain following nerve injury is genetically affected by CACNG2. Genome Res 2010;20:1180-1190.

[25] Singer JB, Hill AE, Nadeau JH Lander ES. Mapping quantitative trait loci for anxiety in chromosome substitution strains of mice. Genetics 2005;169:855-862.

[26] Bennett BJ, Farber CR, Orozco L, et al. A high-resolution association mapping panel for the dissection of complex traits in mice. Genome Res 2010;20:281-290.

[27] Niu N, Qin Y, Fridley BL, et al. Radiation pharmacogenomics: a genome-wide association approach to identify radiation response biomarkers using human lymphoblastoid cell lines. Genome Res 2010;20:1482-1492.

[28] Rodemann HP, Wouters BG. Frontiers in molecular radiation biology/oncology. Radiother Oncol 2011;101:1-6.

[29] Hilbers FS, Boekel NB, van den Broek AJ, et al.Genetic variants in TGFβ-1 and PAI-1 as possible risk factors for cardiovascular disease after radiotherapy for breast cancer. Radiother Oncol 2012;102:115-21.

[30] Pratesi N, Mangoni M, Mancini I, et al. Association between single nucleotide polymorphisms in the XRCC1 and RAD51 genes and clinical radiosensitivity in head and neck cancer. Radiother Oncol 2011;99:356-61.

[31] Barnett GC, Coles CE, Elliott RM, et al. Independent validation of genes and polymorphisms reported to be associated with radiation toxicity: a prospective analysis study. Lancet Oncol 2012;13:65-77.

[32] Andreassen CN, Alsner J. Genetic variants and normal tissue toxicity after radiotherapy: a systematic review. Radiother Oncol 2009;92:299-309.

[33] Franko AJ, Sharplin J, Ward WF Hinz JM. The genetic basis of strain-dependent differences in the early phase of radiation injury in mouse lung. Radiat Res 1991;126:349-356.

[34] Hong JH, Jung SM, Tsao TC, et al. Bronchoalveolar lavage and interstitial cells have different roles in radiation-induced lung injury. Int J Radiat Biol 2003;79:159-167.

[35] Sinke AP, Caputo C, Tsaih SW, et al. Genetic analysis of mouse strains with variable serum sodium concentrations identifies the Nalcn sodium channel as a novel player in osmoregulation. Physiol Genomics 2011;43:265-270.

[36] Keane TM, Goodstadt L, Danecek P, et al. Mouse genomic variation and its effect on phenotypes and gene regulation. Nature 2011;477:289-294.

[37] Freedman ML, Monteiro AN, Gayther SA, et al. Principles for the post-GWAS functional characterization of cancer risk loci. Nat Genet 2011;43:513-518.

[38] Forton JT, Udalova IA, Campino S, Rockett KA, Hull J Kwiatkowski DP. Localization of a long-range cis-regulatory element of IL13 by allelic transcript ratio mapping. Genome Res 2007;17:82-87.

[39] Hallahan DE, Geng L Shyr Y. Effects of intercellular adhesion molecule 1 (ICAM-1) null mutation on radiation-induced pulmonary fibrosis and respiratory insufficiency in mice. J Natl Cancer Inst 2002;94:733-741.

[40] Konigshoff M. Lung cancer in pulmonary fibrosis: tales of epithelial cell plasticity. Respiration 2011;81:353-358.

[41] Cappuccini F, Eldh T, Bruder D, et al. New insights into the molecular pathology of radiation-induced pneumopathy. Radiother Oncol 2011;101:86-92.

[42] Roberts ME, Magowan L, Hall IP Johnson SR. Discoidin domain receptor 1 regulates bronchial epithelial repair and matrix metalloproteinase production. Eur Respir J 2011;37:1482-1493.

[43] Peacock JG, Miller AL, Bradley WD, Rodriguez OC, Webb DJ Koleske AJ. The Abl-related gene tyrosine kinase acts through p190RhoGAP to inhibit actomyosin contractility and regulate focal adhesion dynamics upon adhesion to fibronectin. Mol Biol Cell 2007;18:3860-3872.

[44] Avivi-Green C, Singal M Vogel WF. Discoidin domain receptor 1-deficient mice are resistant to bleomycin-induced lung fibrosis. Am J Respir Crit Care Med 2006;174:420-427.

[45] Guerrot D, Kerroch M, Placier S, et al. Discoidin domain receptor 1 is a major mediator of inflammation and fibrosis in obstructive nephropathy. Am J Pathol 2011;179:83-91.

[46] Barron L Wynn TA. Fibrosis is regulated by Th2 and Th17 responses and by dynamic interactions between fibroblasts and macrophages. Am J Physiol Gastrointest Liver Physiol 2011;300:G723-728.

[47] Yamazaki T, Goya I, Graf D, Craig S, Martin-Orozco N Dong C. A butyrophilin family member critically inhibits T cell activation. J Immunol 2010;185:5907-5914.

[48] Zhong MC Veillette A. Control of T lymphocyte signaling by Ly108, a signaling lymphocytic activation molecule family receptor implicated in autoimmunity. J Biol Chem 2008;283:19255-19264.

[49] Chapman HA. Epithelial-mesenchymal interactions in pulmonary fibrosis. Annu Rev Physiol 2011;73:413-435.

[50] Anscher MS, Thrasher B, Zgonjanin L, et al. Small molecular inhibitor of transforming growth factor-beta protects against development of radiation-induced lung injury. Int J Radiat Oncol Biol Phys 2008;71:829-837.

[51] Yamasaki M, Kang HR, Homer RJ, et al. P21 regulates TGF-beta1-induced pulmonary responses via a TNF-alpha-signaling pathway. Am J Respir Cell Mol Biol 2008;38:346-353.

[52] Shen ZJ, Esnault S, Rosenthal LA, et al. Pin1 regulates TGF-beta1 production by activated human and murine eosinophils and contributes to allergic lung fibrosis. J Clin Invest 2008;118:479-490.

[53] Lin JR, Zeman MK, Chen JY, Yee MC Cimprich KA. SHPRH and HLTF act in a damage-specific manner to coordinate different forms of postreplication repair and prevent mutagenesis. Mol Cell 2011;42:237-249.

[54] Smirnov DA, Morley M, Shin E, Spielman RS Cheung VG. Genetic analysis of radiation-induced changes in human gene expression. Nature 2009;459:587-591.

Table 1. Candidate genes revealed by association analyses

Gene	Chromosome	SM GWAS[1]	SM QWAS[2]	HAPLO QWAS[3]	Interaction	Differential expression (treated vs control)[4]	Differential expression (controls)[5]	SNP location[6]
Cadm1 / Rexo2	9	x	x	x		a, b, c / a, b, c	ba / ba, bc, ca	CNS
Zfp426	9	x				a		3'utr, 5'utr, Cn, i 3, 4, 5, 6
Utrn	10	x				a, c	ba, bc	i 61
Dact1	12	x				c	ca, ba	Cn
5730601F06Rik	9	x				a, b, c		i 2,3,6
Pin1	9	x						i 1, 2
Ubl5	9	x						i 3
Shprh	10	x				a		i 17, 21
Arid4a	12	x						i 23
Kcnip4	5	x						i 1
Prg4	1		x	x				i 1, 4
Tpr	1		x	x	x	a	ba	Cn, i 9
Pi16	17		x			a, c		Cn, i 9, 14
BC003331	1			x				Cn, 3'utr, i 1
Etv6	6			x		a, c		i 1
Zfp30	7			x		a	ba	Cn
Zfp84	7			x		a		3'utr, 5'utr, i 1
Cdkn1a	17			x		a, b, c		Cn, 3'utr, i 1
Abl2	1				x		ba, bc	3'utr
Slamf6	1				x	b, c	ba, bc	i 1
Myo1f	17				x	c	ba, bc	i 1, downstream
Btnl1	17				x			downstream
Notch4	17				x	a, b, c		i 17, upstream
Ddr1	17				x	c		i 7
Btbd9	17				x	a, c	ca, ba, bc	i 4
Mrps18b	17				x	a, c		i 1

[1] Single Marker Genome Wide Association Analysis
[2] Single Marker QTL Wide Association Analysis
[3] Haplotype Analysis QTL Wide
[4] a = differentially expressed in A/J Treated vs A/J Control; b = differentially expressed in C57BL6/J Treated vs C57BL6/J Control; c = differentially expressed in C3H/HeJ Treated vs C3H/HeJ Control. Treated is defined as irradiated with a single dose of 18Gy to the thorax (ref 6).
[5] ab = differentially expressed in A/J vs C57BL6/J; bc = differentially expressed in C57BL6/J vs C3H/HeJ; ca = differentially expressed in C3H/HeJ vs A/J
[6] CNS=Conserved Noncoding Sequence, Cn=Coding nonsynonymous SNP, i=intron

Figure Legends

Figure 1. Radiation-induced lung phenotype of consomic B6.A mice. Mice of each strain were exposed to 18 Gy whole thorax irradiation and euthanized when moribund or at 26 weeks post treatment. (a) percent fibrotic lung tissue in Trichrome stained histological sections (b) alveolitis score derived from semi quantitative evaluation of histological sections and (c) time post treatment to develop respiratory distress. Phenotypes are presented as the mean ± std error of groups of 7-15 mice. The average phenotypes of untreated control mice were an alveolitis score of 1.8 and 0% fibrosis. *indicates a significant difference in phenotype between the parental strains, $p<0.05$; + indicates a significant difference in phenotype compared to the A/J strain, $p<0.05$; #indicates a significant difference in phenotype compared to the B6 strain, $p<0.05$.

Figure 2. Murine strain difference in lung response to radiation. Mice of each strain were exposed to 18 Gy whole thorax irradiation and euthanized when moribund, or at 26 weeks post treatment which was the end of experiment (EOE). (a) images of Masson's trichrome-stained lung sections from strains indicating different fibrosis responses to whole thorax irradiation; magnification = 200X. (b) percent fibrotic lung tissue in Trichrome stained histological sections (c) alveolitis score derived from semi quantitative evaluation of histological sections and (d) time post treatment to develop respiratory distress. Phenotypes are presented as the mean ± std error of groups of 9-20 mice. The average phenotypes of untreated control mice were an alveolitis score of 1.6 and 0% fibrosis.

Figure 3. Murine strain difference in inflammatory lung response to radiation. Mice of each strain were exposed to 18 Gy whole thorax irradiation and euthanized when moribund. (a)

mast cell count from Toluidine blue stained lung sections post radiation treatment. Mast cell counts in control mice were 0.07±0.05 cells/mm^2 and (b) cell types in bronchoalveolar lavage (BAL). Phenotypes are presented as the mean ± SEM of groups of 9-20 mice.

Figure 4. GWA and QWA of radiation-induced pulmonary fibrosis in mice. The fibrosis phenotype data of Figure 2 were analysed with (a) genome-wide and (b) chromosome 9 QTL SNP genotypes and p values for association computed using EMMA. The dashed grey line indicates the 0.05 significance threshold which for the genome-wide dataset corresponds to a point-wise p-value of $2.68x10^{-6}$.

Figure 1

Click here to download high resolution image

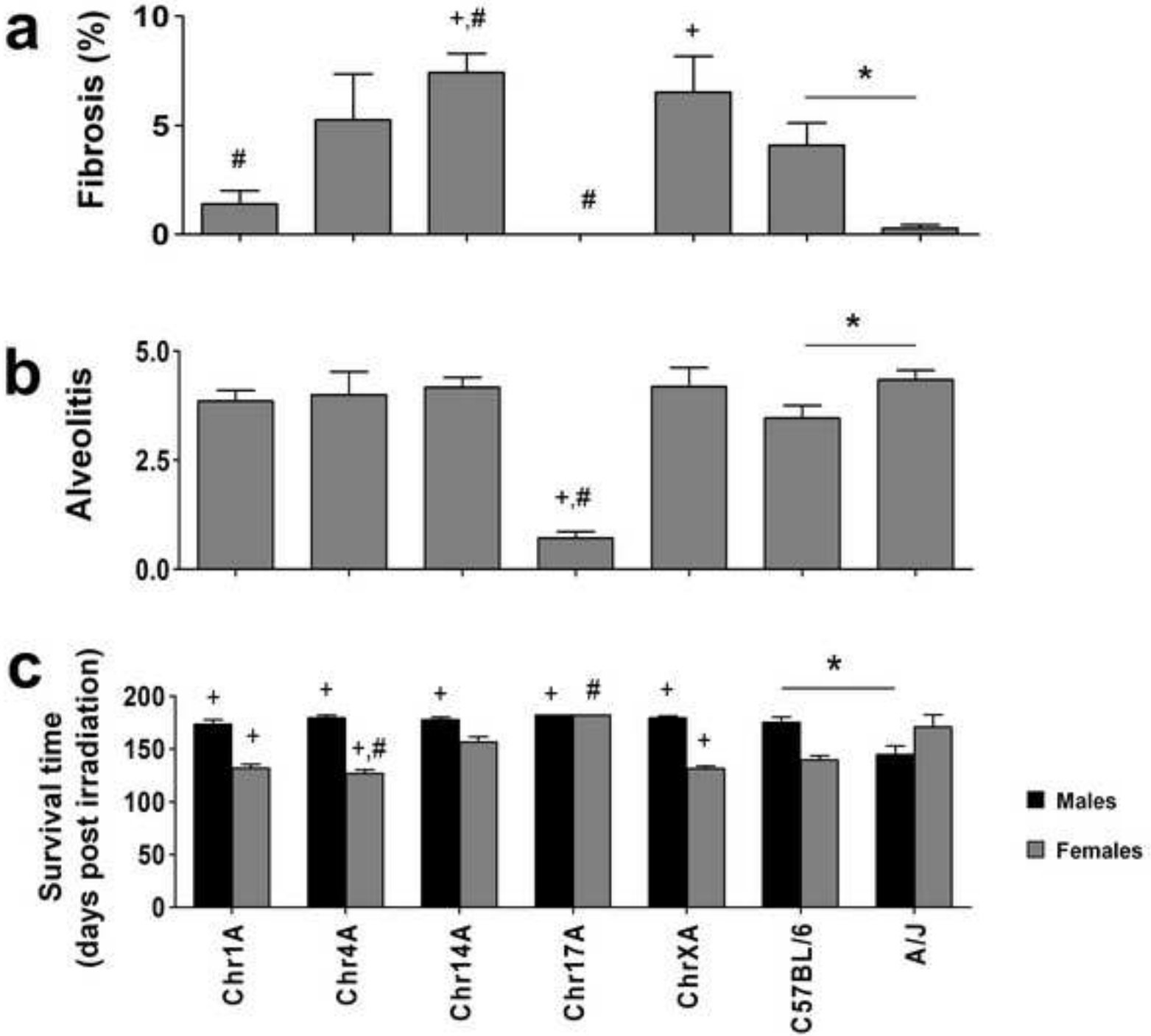


Figure 2

Click here to download high resolution image

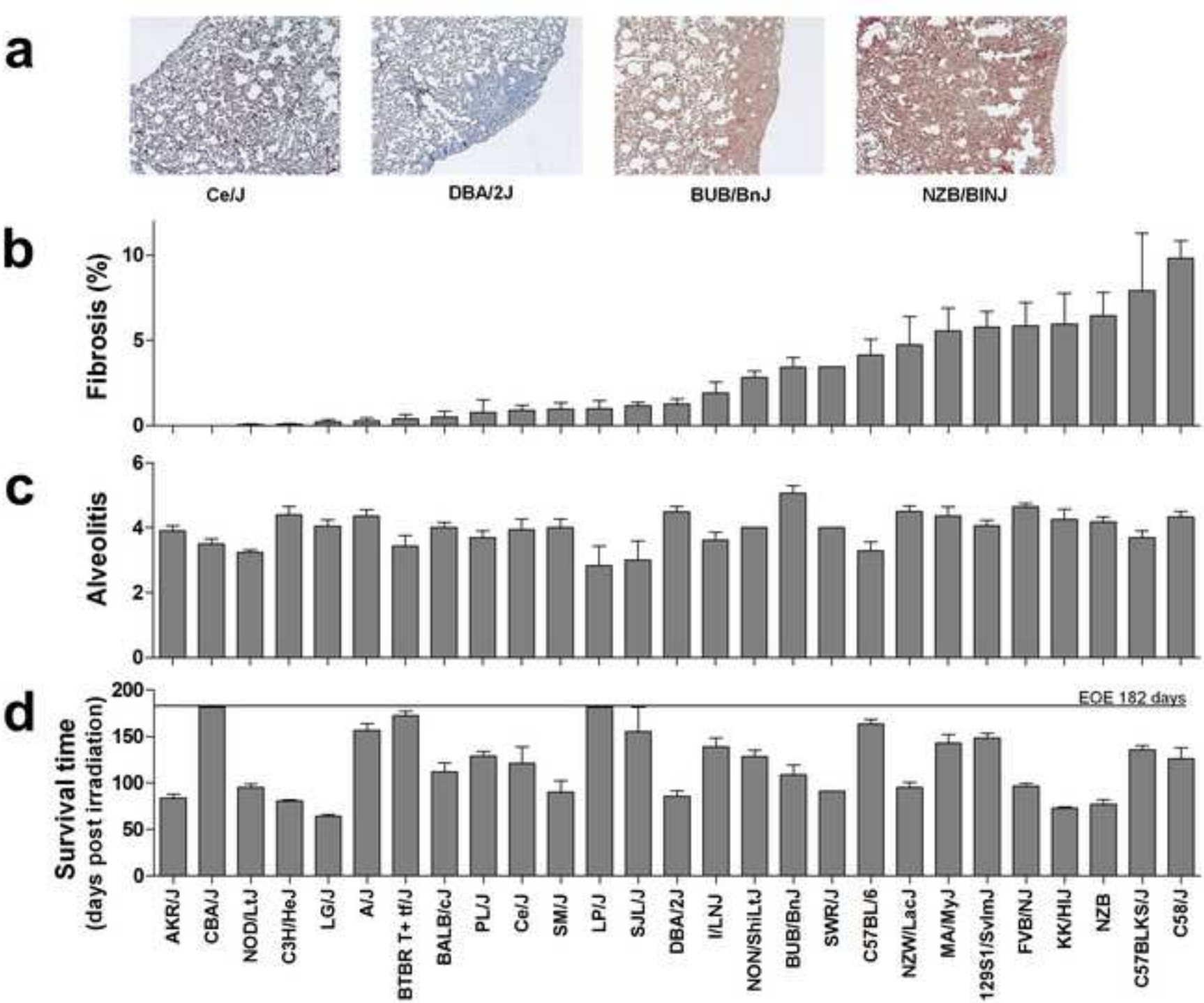


Figure 3

Click here to download high resolution image

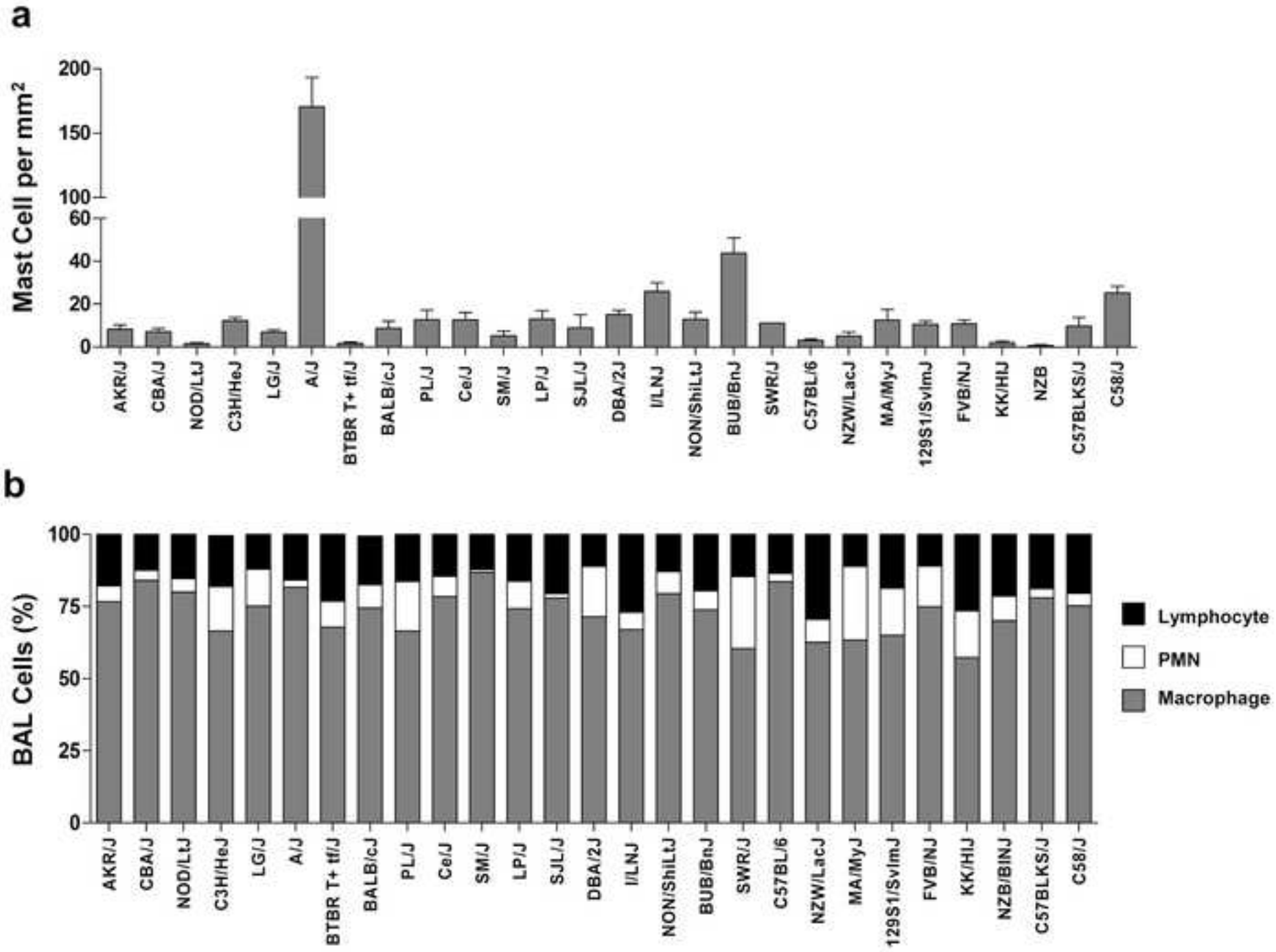
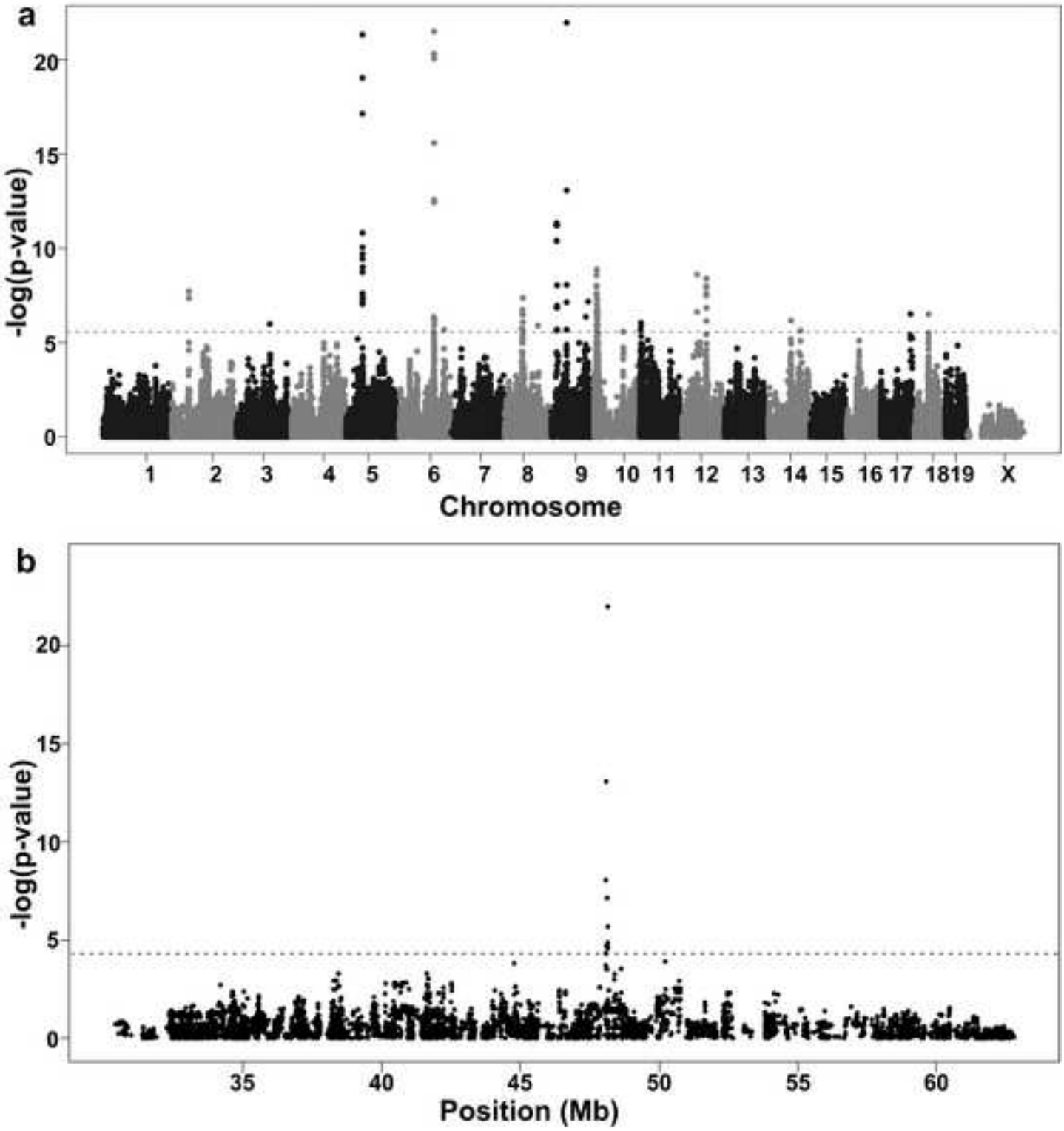


Figure 4

Click here to download high resolution image



SUPPLEMENTARY DATA

Materials and Methods

Mice: Mice of inbred (AKR/J, CBA/J, NOD/ShiLtJ, C3H/HeJ, LG/J, A/J, BTBR T+ tf/J, BALB/cJ, PL/J, CE/J, SM/J, LP/J, SJL/J, DBA/2J, I/LnJ, NON/ShiLtJ, BUB/BnJ, SWR/J, C57BL/6J, NZW/LacJ, MA/MyJ, 129S1/SvImJ, FVB/NJ, KK/HlJ, NZB/BlNJ, C57BLKS/J, C58/J) and consomic B6.A strains (C57BL/6J-Chr $1^{A/J}$/NaJ, C57BL/6J-Chr $4^{A/J}$/NaJ, C57BL/6J-Chr $14^{A/J}$/NaJ, C57BL/6J-Chr $17^{A/J}$/NaJ, C57BL/6J-Chr $X^{A/J}$/NaJ) were purchased from the Jackson Laboratory (Bar Harbor, USA) and housed in the animal facility of the Meakins-Christie Laboratories. All mice were handled according to guidelines and regulations of the Canadian Council on Animal Care.

Radiation treatment: Mice were treated at 8 weeks of age. Lung damage was elicited by whole thorax radiation exposure (18 Gy; dose rate 0.7 Gy/minute) using a Gamma cell Cesium-137 unit as previously described [1-5]. The irradiated mice were sacrificed when moribund, or at 26 weeks after treatment. The mice were weighed weekly from eight weeks after radiation and mice losing > 20% body weight, and exhibiting distress through ruffled fur, accelerated breathing and hunched posture, were sacrificed as described previously [1]. The control mice (five of each of B6, C3H, FVB/NJ, & A/J strains, four NOD/ShiLtJ, seven BTBR T+ tf/J, four C57BL/6J-Chr $1^{A/J}$/NaJ and three C57BL/6J-Chr $4^{A/J}$/NaJ mice) were not treated and were euthanized at the 15-26 week time points.

Histology: At necropsy, bronchoalveolar lavage (BAL) was performed by cannulating the trachea and retrieving cells from three 1-mL injections of phosphate buffered saline. The lungs

were then removed and the single left lobe of each mouse was perfused with 10% neutral buffered formalin and submitted for histological processing. Lung sections were stained with Masson's Trichrome and the area of fibrosis in the left lung lobe was determined from a user drawn region surrounding the fibrosis (Image Pro Plus Software) compared to the area of the entire lobe to yield the percent of pulmonary fibrosis for individual mice [1,3]. To assess alveolitis, hematoxylin and eosin (H&E) stained left lung sections were evaluated through semi-quantitative histology [1,3]. Mast cell numbers were determined by counting the positive cells present in 10 fields of a lung histological section stained with Toluidine blue, at a 400x magnification. All scoring was completed by a user blinded to mouse strain and treatment.

Bronchoalveolar Lavage Fluid (BAL) Analysis: The BAL fluid was centrifuged (302g for 10 minutes @ 4°C) and the supernatant was removed and stored at -85°C. The cellular pellet was re-suspended in 0.25mL PBS. Inflammatory cell counts were performed (400x magnification) on cytocentrifuged cells (214.2g for three minutes), after staining with a hematoxylin-eosin kit (Hema-3 Stain Set by Protocol) and are reported as the percentage of 500 counted cells. To assess correlations among disease phenotypes pairwise Pearson's correlation coefficients and corresponding p-values were calculated using the rcor.test function in R (http://cran.r-project.org/).

Genome-wide association studies (GWAS): GWAS were completed using 236 mice from 27 inbred strains and the following phenotypes: percent fibrotic lung tissue, survival post irradiation, percent of lymphocytes & neutrophils in BAL and lung tissue mast cell count. The

phenotype data include the response of one SWR/J mouse, which was included as the measured phenotype agrees with that previously reported for this strain [6].

Genotype data consisted of the CGD1 (http://cgd.jax.org/index.php) imputed dataset downloaded from the Mouse Phenome Database (Jackson Laboratory, Bar Harbor, ME). This dataset was filtered for SNPs imputed with confidence scores >0.9 as in [7], and SNPs with fewer than 20 strains typed/imputed or with a minor allele frequency <3/27 were removed as has been used by others [7,8]. The 1,885,108 remaining SNPs span the mouse genome at an average of 1 per 1.8 kB and were assessed for association through single marker analysis and haplotype analysis.

In single marker analysis genotype-phenotype associations were evaluated for each individual SNP using efficient mixed model association (EMMA, ref [9]), a method which takes into account the relatedness between the inbred strains as used by others [7,10,11]. For each SNP a two-sided p-value was obtained by testing the null hypothesis of no association between genotype and phenotype, using the R implementation of the EMMA package (http://mouse.cs.ucla.edu/emma). A locus was considered to be significantly associated with the phenotype if its SNP score (-log10(pvalue)) was above the genome-wide threshold and it contained at least two SNPs within 50 kB of each other [12]. The significance thresholds were determined through permutation testing as in [8,11]. In detail, for each of 1000 permutations, phenotype values were randomly shuffled, genotype-phenotype associations tested and the minimal nominal p-value retained. The resulting 1000 minimal p-values were sorted in ascending order and the 5% quantile of the distribution was taken as the genome-wide significance threshold (global p-value of 0.05) as used by Lu *et al.* [8].

QTL-wide association studies (QWAS): To specifically assess the SNP variation within the previously identified linkage intervals [1,13] for association with radiation-induced lung disease we performed QWAS. For this SNPs located in linkage regions on chromosomes 1,6,7,9 & 17 and the percent fibrosis phenotype data of the GWAS were re-analysed with the single marker method. We added the condition that the A/J and C3H alleles be different from the B6 allele for the regions on chromosomes 1 & 17; and that the C3H allele be different from the B6 allele for the regions on chromosomes 6, 7 & 9. The significance level for association was determined through permutation testing with the QTL region SNPs and the 5^{th} percentile of the resultant distribution was taken as the QTL-wide significance threshold.

Linkage interval genotypes were also tested for association with the fibrosis phenotype using haplotype blocks derived from the SNP data through a build-and-step-back approach, based on the method developed by Cahan *et al.* [14]. Our algorithm iteratively joins consecutive SNPs into haplotype blocks by clustering strains based on their genotypes. A consensus haplotype (a vector composed of the most frequent genotypes at each string position) was derived for each cluster of strains and we allowed an overall error rate of 10% of the block length between the genotypes of each individual strain and the consensus haplotype (e.g in a haplotype block of 100 SNPs shared among 4 strains, we allowed a total of 10 SNPs to be different among the 4 individual strain genotypes and the consensus haplotype). At each step in the algorithm the haplotype was "built" by adding the consecutive SNP, if the above condition was met. When this condition was not met our algorithm "stepped back" and re-analyzed the shorter block obtained by iteratively subtracting the first SNP in the current block until a valid haplotype was obtained. Thus the method generated larger, overlapping haplotypes, and not adjacent blocks, as in Cahan *et al.* To ensure a maximum number of polymorphisms were evaluated in this approach

we used the complete SNP dataset without applying the minor allele frequency and genotyping rate filters and thus only SNPs with a minor allele frequency of zero in the 27 strains were removed. The "step-back" approach greatly increased the computational time, therefore the analysis was restricted to the QTL level.

In this analysis each haplotype block was represented by a string of "labels" for each strain indicating the haplotype belonging to each strain. The association of the phenotype (percent fibrosis) with the haplotype "labels" was assessed using an analysis of variance (ANOVA). The weighted bootstrap significance threshold was too stringent ($<10^{-32}$) and yielded no significant results. We therefore present the ten most significant haplotypes in each QTL as in (12), for which the lowest nominal p-values ranged from 10^{-12} for chr 7 to 10^{-19} for chr 9.

SNP-SNP interaction between Radpf1 and Radpf2: To determine whether an interaction of genotypes within the linkage intervals *Radpf1* on chromosome 17 and *Radpf2* on chromosome 1 contributes to the fibrosis phenotype, as suggested in [13], we tested the association between the phenotype and all possible pairwise sets of SNPs as predictors through a linear model including a SNP-SNP interaction term. The significance of the interaction term was determined using an F test by comparing a full model including the interaction term with the partial model that excludes the interaction variable. Approximately 8 million tests were performed and the results were corrected for multiple testing using the Benjamini-Hochberg FDR approach.

Candidate Gene Identification: Genes with SNP variation meeting the GWAS/QWAS significance criteria were initially defined using BioMart feature of Ensembl which incorporated NCBI build 37 (http://www.ensembl.org).

These genes were next filtered to reveal those that are expressed in lung, which was

defined as genes having a “Present” call as obtained using the MAS5.0 algorithm, in a prior microarray study of radiation induced lung disease [4]. The lung expression of the “Present” genes was confirmed using data from two additional microarray experiments (GEO database - GDS1492 (ref [15]) and GDS1649 (ref [16]). Similarly, genes not expressed in the lung were deemed to be absent in the two other microarray studies. The only exception was *Kcnip4* which shows lung expression in the majority of assays in the GEO database, including GDS1492, and was therefore retained in the list of candidates. Of the genes filtered for pulmonary expression, further evaluation revealed those harbouring SNPs which create a non synonymous mutation, or interrupt a splice site as in [7,17,18].

SNPs of significant association were retained in the candidate list if determined to map to conserved non coding regions (CNS) of the genome. In this analysis, for each SNP region an alignment was performed between the mouse sequence and all vertebrate genomes available in the database Genome Browser (http:// genome.ucsc.edu) using the MULTIZ method [19]. The most conserved elements were identified and scored by PhastCons [20] and were considered as potential regulatory sequences if they were not located in gene exons, their length >100 bp and the PhastCons LOD score was higher than 70, as defined by others [21,22].

The QWAS-derived candidate genes with significantly associated SNPs in potential regulatory regions (intronic, 5’ or 3’ UTR, CNS) were further filtered for those with strain dependent pulmonary differential expression (in the untreated condition or post thoracic radiation), using our gene expression profiles [4] as in [7,10].

Reference List

[1] Haston CK, Begin M, Dorion G Cory SM. Distinct loci influence radiation-induced alveolitis from fibrosing alveolitis in the mouse. Cancer Res 2007;67:10796-10803.

[2] O'Brien TJ, Letuve S Haston CK. Radiation-induced strain differences in mouse alveolar inflammatory cell apoptosis. Can J Physiol Pharmacol 2005;83:117-122.

[3] Lemay AM Haston CK. Radiation-induced lung response of AcB/BcA recombinant congenic mice. Radiat Res 2008;170:299-306.

[4] Paun A, Lemay AM Haston CK. Gene expression profiling distinguishes radiation-induced fibrosing alveolitis from alveolitis in mice. Radiat Res 2010;173:512-521.

[5] Thomas DM, Fox J Haston CK. Imatinib therapy reduces radiation-induced pulmonary mast cell influx and delays lung disease in the mouse. Int J Radiat Biol 2010;86:436-444.

[6] Sharplin J Franko AJ. A quantitative histological study of strain-dependent differences in the effects of irradiation on mouse lung during the intermediate and late phases. Radiat Res 1989;119:15-31.

[7] Liu PY, Vikis H, James M, et al. Identification of Las2, a major modifier gene affecting the Pas1 mouse lung tumor susceptibility locus. Cancer Res 2009;69:6290-6298.

[8] Lu Y, Liu P, James M, et al. Genetic variants cis-regulating Xrn2 expression contribute to the risk of spontaneous lung tumor. Oncogene 2010;29:1041-1049.

[9] Kang HM, Zaitlen NA, Wade CM, et al. Efficient control of population structure in model organism association mapping. Genetics 2008;178:1709-1723.

[10] Yang H, Wang JR, Didion JP, et al. Subspecific origin and haplotype diversity in the laboratory mouse. Nat Genet 2011;43:648-655.

[11] Bennett BJ, Farber CR, Orozco L, et al. A high-resolution association mapping panel for the dissection of complex traits in mice. Genome Res 2010;20:281-290.

[12] Niu N, Qin Y, Fridley BL, et al. Radiation pharmacogenomics: a genome-wide association approach to identify radiation response biomarkers using human lymphoblastoid cell lines. Genome Res 2010;20:1482-1492.

[13] Haston CK, Zhou X, Gumbiner-Russo L, et al. Universal and radiation-specific loci influence murine susceptibility to radiation-induced pulmonary fibrosis. Cancer Res 2002;62:3782-3788.

[14] Cahan P Graubert TA. Integrated genomics of susceptibility to alkylator-induced leukemia in mice. BMC Genomics 2010;11:638.

[15] Haston CK, Tomko TG, Godin N, Kerckhoff L Hallett MT. Murine candidate bleomycin induced pulmonary fibrosis susceptibility genes identified by gene expression and sequence analysis of linkage regions. J Med Genet 2005;42:464-473.

[16] Stearman RS, Dwyer-Nield L, Zerbe L, et al. Analysis of orthologous gene expression between human pulmonary adenocarcinoma and a carcinogen-induced murine model. Am J Pathol 2005;167:1763-1775.

[17] Harrill AH, Watkins PB, Su S, et al. Mouse population-guided resequencing reveals that variants in CD44 contribute to acetaminophen-induced liver injury in humans. Genome Res 2009;19:1507-1515.

[18] Sudbery I, Stalker J, Simpson JT, et al. Deep short-read sequencing of chromosome 17 from the mouse strains A/J and CAST/Ei identifies significant germline variation and candidate genes that regulate liver triglyceride levels. Genome Biol 2009;10:R112.

[19] Blanchette M, Kent WJ, Riemer C, et al. Aligning multiple genomic sequences with the threaded blockset aligner. Genome Res 2004;14:708-715.

[20] Siepel A, Bejerano G, Pedersen JS, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res 2005;15:1034-1050.

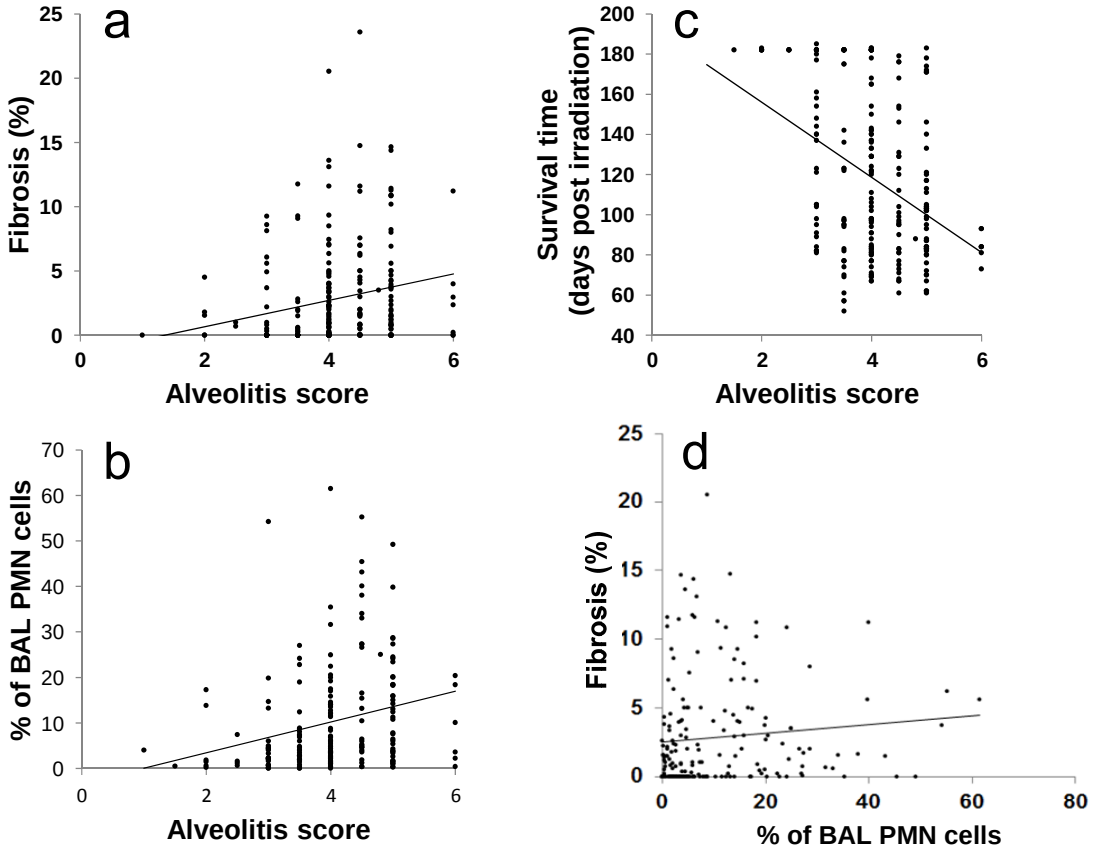
[21] Loots GG, Locksley RM, Blankespoor CM, et al. Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. Science 2000;288:136-140.

[22] Li Q, Harju S Peterson KR. Locus control regions: coming of age at a decade plus. Trends Genet 1999;15:403-408.

Supplementary Table 1. Significantly interacting SNP pairs in *Radpf1* and *Radpf2*

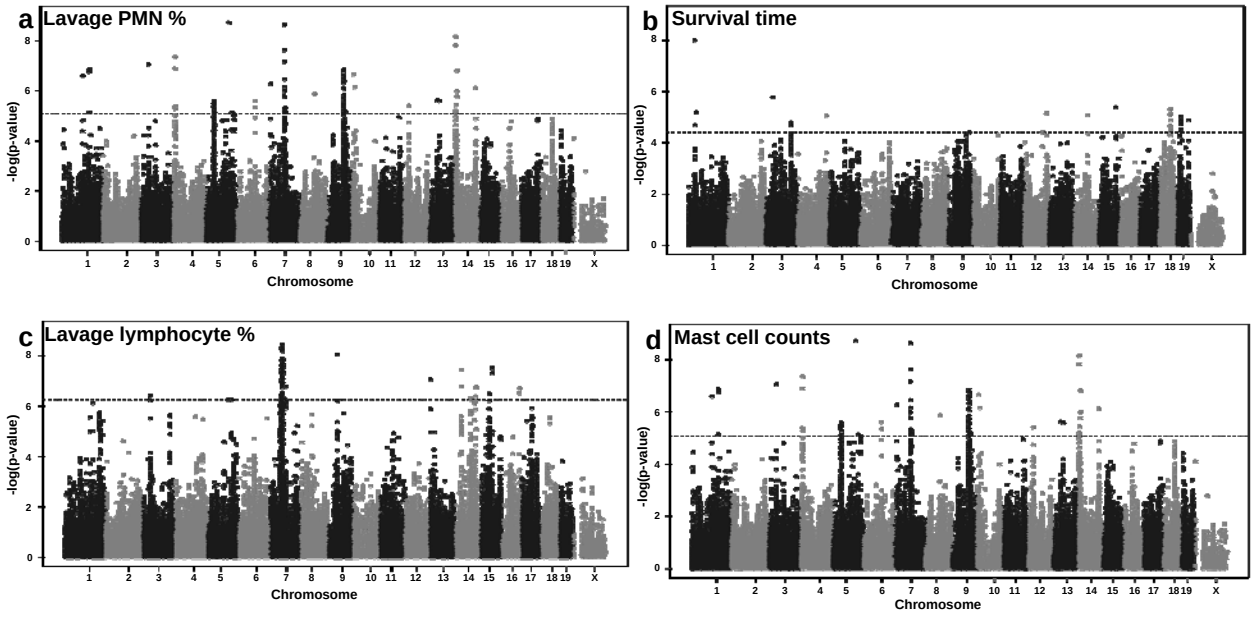
***Radpf1* SNP**	***Radpf2* SNP**	**p-value**	**FDR**	***Radpf1* nearest gene**	***Radpf2* nearest gene**
rs50253706	rs31168022	2.67E-09	0.01	*Ddr1*	*Tpr* intron 9
rs51899510	rs31168022	2.60E-09	0.01	*Ddr1*	Tpr intron 9
rs33801987	rs31168022	1.01E-08	0.025	*Btbd9* intron 4	*Tpr* intron 9
rs46654461	rs31272943	1.75E-07	0.134	*Myo1f*	*Abl2* 3'UTR
rs49838267	rs31272943	1.96E-07	0.134	*Myo1f* intron 1	*Abl2* 3'UTR
rs33444918	rs31529220	1.72E-07	0.134	*Btnl1*	*Slamf6* intron 1
rs33444918	rs31528128	2.66E-07	0.134	*Btnl1*	*Slamf6* intron 1
rs47182407	rs31168022	2.35E-07	0.134	*Mrps18b* intron 1	*Tpr* intron 9
rs51682784	rs31168022	2.03E-07	0.134	*Ddr1* intron 7	*Tpr* intron 9
rs29518121	rs46915727	2.17E-07	0.134	*Notch4* intron 17	intergenic

A linear model was fitted to the phenotypic data with *Radpf1* and *Radpf2* SNPs as covariates. The significance of the interaction term between each pair of SNPs was determined by comparing the model with interaction versus a simple additive model. The first 10 significant interacting SNP pairs are shown together with the false discovery rates and the gene containing or located near the SNP.



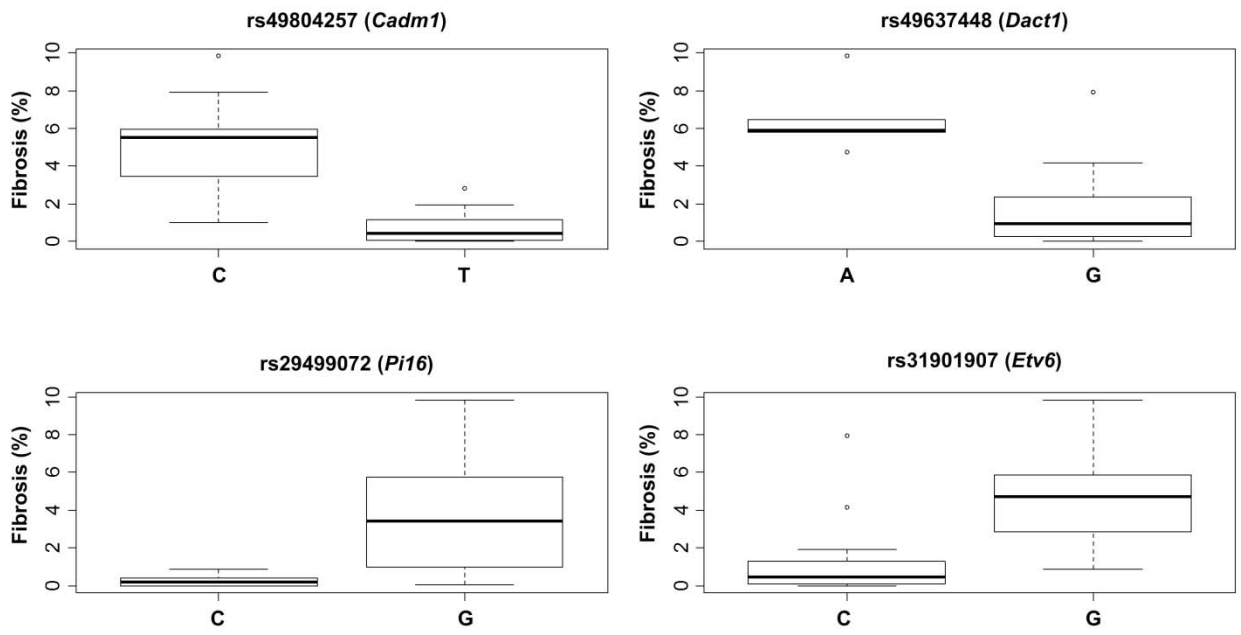
Supplementary Figure 1. Correlations between radiation-induced lung response phenotypes in the panel of inbred strains.

Mice were exposed to 18 Gy whole thorax irradiation and euthanized when moribund or at 26 weeks post treatment. Phenotypes as presented in Figure 2. Correlations of alveolitis score and (a) fibrosis score, (b) bronchoalveolar lavage neutrophil percent (c) survival time post radiation treatment; and of (d) the fibrosis score to bronchoalveolar lavage neutrophil percent.

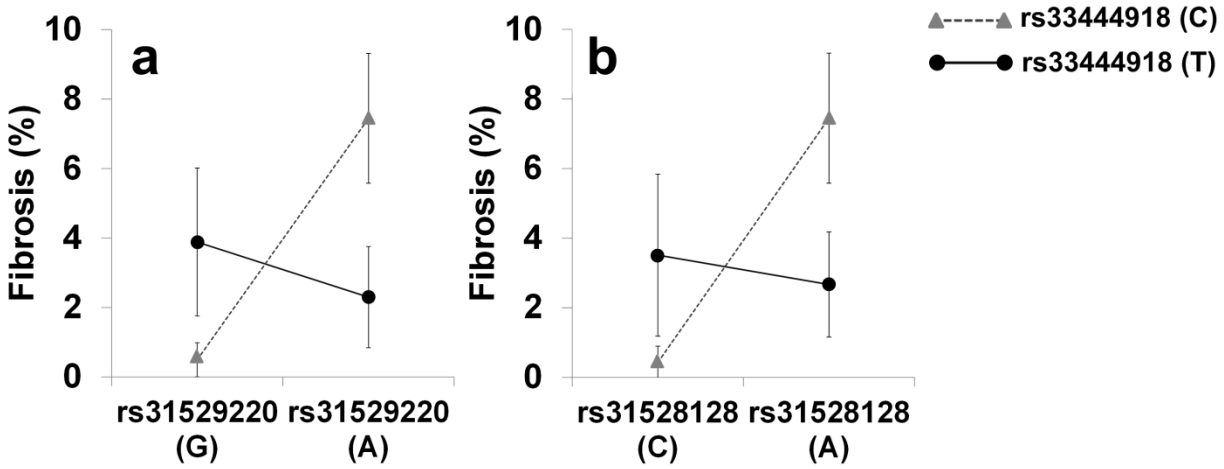


Supplementary Figure 2. GWA of radiation-induced lung phenotypes in mice.

The survival phenotype data of Figure 1 and the inflammatory response data of Supplementary Figure 2 were analysed with genome-wide SNP genotypes and p values for association computed using EMMA. (a) Lavage PMN percent, (b) Survival time post radiation, (c) lavage lymphocyte percent, (d) mast cell counts. The horizontal lines indicate the 0.05 genome-wide significance thresholds.



Supplemental Figure 3. Genotype-phenotype associations of radiation-induced pulmonary fibrosis in mice. Each boxplot shows the median, lower and upper quartiles and extreme values of the fibrosis levels of the inbred strains, as presented in Figure 2, according to their genotype at a particular SNP of significant association. The polymorphism as well as the gene it is located in is indicated in the title of each plot.



Supplemental Figure 4. Epistatic interaction between SNPs in Radpf1 and Radpf2 influences radiation-induced pulmonary fibrosis in mice. Linear models were adjusted for the phenotype of radiation-induced pulmonary fibrosis in 27 inbred strains using the two SNPs as covariates. The significance of the interaction term was obtained by comparing the complete model (with interaction) versus the simple additive model (no interaction). The x axis shows the genotypes for A) rs31529220 B) rs31528128 on chromosome 1. The genotypes for the interacting SNP, rs33444918 on chromosome 17, are shown in the legend. Phenotype values on the y axis are presented as mean ± stdev for each genotype group.