Haplotype association mapping of acute lung injury in mice implicates activin A receptor, type 1

Supplemental Material

George D Leikauf, Vincent J Concel, Pengyuan Liu, Kiflai Bein, Annerose Berndt, Koustav Ganguly, An Soo Jang, Kelly A Brant, Maggie Dietsch, Hannah Pope-Varsalona, Richard A Dopico, Jr., YP Peter Di, Qian Li, Louis J Vuga, Mario Medvedovic, Naftali Kaminski, Ming You, Daniel R Prows

Supplemental Material

Methods

Animal Exposure: This study was performed in accordance with the Institutional Animal Care and Use Committee of the University of Pittsburgh (Pittsburgh, PA) and mice were housed under specific pathogen free conditions. Forty inbred mouse strains were used in the haplotype association mapping (n = 365) and SM/J (n = 71) and 129X1/SvJ (n = 71) mouse strains were used for additional characterization of lung injury (6-8 weeks female; The Jackson Laboratory, Bar Harbor, ME) . Animals were exposed to filtered air (control) or acrolein (10 ppm, 24 h) generated and monitored as described previously (8), and survival time was recorded.

To examine acrolein-induced changes in lung histology, bronchoalveolar lavage, and lung transcripts, 129X1/SvJ and SM/J mice were exposed to filtered air (0 h, control) or acrolein (10 ppm for 6, 12 or 17h). Mice were killed immediately after acrolein exposure by intraperitoneal injection of pentobarbital sodium (100 mg/kg; Nembutal, Abbott Laboratories, Chicago, IL) and severing of the posterior abdominal aorta. To obtain tissue for mRNA analysis (n = 7-8 mice/strain), the diaphragm was punctured, and the chest cavity opened. Lungs were excised, frozen in liquid nitrogen, and stored (-70°C). To obtain tissue for histology (n = 3 mice/strain), the chest wall was left intact, a cannula was inserted into the trachea and the lung was instilled with phosphate-buffered saline containing 3.7% formaldehyde (pressure: 28 cm H_2O, Cat. No. SF100-4, Thermo Fisher Scientific, Pittsburgh, PA). The trachea was ligated, lungs removed, and the inflated lung was immersed in fixative (24 h, 4°C). Fixed lungs were washed with Dulbecco's Phosphate-Buffered Saline containing Ca^{2+}, Mg^{2+}, 6.1 mM D-glucose, and 0.33 mM sodium pyruvate (DPBS; Cat. No. 14287–080, Life Technologies, Carlsbad, CA), dehydrated through a series of graded ethanol solutions (30-70%), and processed in paraffin blocks (Hypercenter XP, Shandon, Ramsey, MN). The lung tissue was sectioned (5 μm) and stained with hematoxylin and eosin.

To obtain bronchoalveolar lavage fluid (n = 5 mice/strain), a cannula was inserted into the trachea and the lung was instilled with Ca^{2+}, Mg^{2+}-free Hank's Balanced Salt Solution (Cat. No. 14175-095, Life Technologies) and the recovered lavage fluid placed on ice until protein analysis was performed. Total protein in cell-free supernatants was measured using a bicinchoninic acid assay (BCA; Cat. No. 23325, Thermo Scientific) using bovine serum albumin (BSA) as a standard. Nitrite concentration was determined was analyzed using a fluorometric method in which nitrite reacted with 2,3-diaminonaphthalene to produce a fluorescent compound (excitation 365 nm and emission 450 nm).

Microarray: RNA quantity was initially assessed by spectrophotometer (Nanodrop ND-1000 Thermo Scientific, Wilmington, DE) and quality was assessed by electrophoresis (Agilent 2100 Bioanalyzer, Agilent Technologies, Santa Clara,, CA). RNA (0.5 μg) was cyanine 3-labeled and cDNA synthesized (RNA Spike In – One Color, 5188-5282, Agilent Technologies). Labeled cRNA was

transcribed from cDNA (Quick-Amp Labeling Kit - One Color, 5190-0442, Agilent Technologies). The labeled cRNA was quantified (Nanodrop ND-1000 and Agilent 2100 Bioanalyzer), and hybridized (65°C, 17 h) (Gene Hybridization, 5188-5242, Agilent Technologies) onto the array (Whole Mouse Genome Kit 4x44K, G4122F, Agilent Technologies). Arrays were washed and scanned (DNA Microarray Scanner, G2505C, Feature Extraction Version 10.7.3.1, Agilent Technologies). Five microarrays were obtained for each strain (SM/J or 129X1/SvJ) and each time (0, 6, or 12 h) to yield a total of 30 microarrays.

Statistical analysis was performed using R statistical software and the limma Bioconductor package (S1). Data normalization was performed in two steps for each microarray. First, background adjusted intensities were log-transformed and the differences (M) and averages (A) of log-transformed values were calculated as $M = \log2(X1) - \log2(X2)$ and $A = [\log2(X1) + \log2(X2)] / 2$, where X1 and X2 denote the Cy5 and Cy3 intensities, respectively. Second, normalization was performed by fitting the array-specific local regression model of M as a function of A. Normalized log-intensities for the two channels were then calculated by adding half of the normalized ratio to A for the Cy5 channel and subtracting half of the normalized ratio from A for the Cy3 channel. The statistical analysis was performed for each gene separately by fitting the following Analysis of Variance model: $Y_{ijk} = \mu + A_i + S_j + Ck + \varepsilon_{ijk}$, where Y_{ijk} corresponds to the normalized log-intensity on the ith array, with the jth treatment, and labeled with the kth dye (k = 1 for Cy5, and 2 for Cy3). μ is the overall mean log-intensity, A_i is the effect of the ith array, S_j is the effect of the jth treatment and C_k is the gene-specific effect of the kth dye. Estimated log 2 fold changes were calculated from the ANOVA models, and resulting t-statistics from each comparison were modified using an intensity-based empirical Bayes method (IBMT) (S2).

Ingenuity Pathway Analysis: Pathway enrichment was determined using significant values (>2-fold, and $p<0.01$ by ANOVA as compared to strain-matched control) analyzed using the Pathway/ Function/List features of Ingenuity Pathways Analysis (Ingenuity® Systems, www.ingenuity.com). For the SM/J mouse strain the number of elements analyze were: increased (6 h): 599, increased (12 h): 1191, decreased (6 h): 313, decreased (12 h): 1050. For the SM/J mouse strain the Pathway/Function/List eligible values were: increased (6 h): 350, increased (12 h): 689, decreased (6 h): 178, decreased (12 h): 591 (**Supplemental Table S1**). For the 129X1/SvJ mouse strain the number of elements analyze were: increased (6 h): 608, increased (12 h): 933, decreased (6 h): 418, decreased (12 h): 800. For the 129X1/SvJ mouse strain the Pathway/Function/List eligible values were: increased (6 h): 352, increased (12 h): 517, decreased (6 h): 258, decreased (12 h): 479. The top 3 enriched canonical pathway, biological function, or toxicology list were selected based on the combined 6 and 12 h -log (P).

Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR): To determine transcript levels of genes previously associated with acute lung injury and to contrast genes identified, quantitative Real-

Time Polymerase Chain Reaction (qRT-PCR) was used. Total RNA was extracted from the lung with TRIreagent (Cat No. 93289, Fluka, St Louis, MO), and quality and quantity were assessed by A260/A280 absorbance (NanoDrop 1000, Thermo Scientific, Wilmington, DE). For qRT-PCR, 200 ng RNA was reverse transcribed into first-strand cDNA using a high capacity cDNA archive kit [Applied Biosytems Inc. (ABI), Foster City, CA] in a 20 μl reaction volume. An aliquot of the cDNA synthesis product (2 μl) was used in a subsequent qRT-PCR reaction using 10 μl universal PCR master mix (TaqMan Universal PCR Master Mix, ABI), 1.0 □l primer mixture, and 7 μl RNAse-free water. The primers used were obtained from ABI. The qRT-PCR analysis was performed with an Applied Biosystems 7900HT System using the following conditions: 95°C for 12 min followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. The comparative cycle number threshold (CT) method (ΔΔCT) was used to quantify the transcript expression levels. The change in CT value ($\Delta CT = CT^{\text{gene of interest}} - CT^{RPL32}$) was calculated for each sample. The comparative ΔΔCT calculation involved finding the difference between each sample's ΔCT and the mean ΔCT for the control samples. These values were transformed to absolute values using the formula: comparative level (log 2 fold change) = $2^{-\Delta\Delta CT}$

In silico protein analysis: ExPASy PROSITE scanning was performed to detect the location and effect of each amino acid polymorphism in relation to the known functional domains of a protein as described previously (Sigrist 2002, Ganguly 2007).

Genotyping of mouse SNP in *Acvr1* 5'UTR promoter: Genomic DNA was prepared from tail clips of 28 mouse strains (Wizard Genomic DNA Purification Kit, Cat. No. A1120, Promega, Madison, WI). To genotype the Acvr1 promoter, the region spanning nucleotides 58,369,109 to 58,368,414 (NCBI Reference Sequence NC_000068.8) was PCR-amplified using the primers:

forward – 5'-AGGGAGGGAGACTGACTTCATCTTATCATCC- 3'

reverse – 5'-TCAACGTGGCAGAGAACTTACCCTGTAG- 3'.

The 696 bp PCR product was gel purified (Qiaquick, Qiagen, CA) and the nucleotide sequence determined by automated sequencing (University of Pittsburgh, Proteomics and Genomics Core Laboratory, Pittsburgh, PA). To identify SNP polymorphisms, the sequences were aligned using ClustalW multiple sequence alignment program (SDSC Biology Workbench, University of California, San Diego, CA).

Electrophoretic mobility shift assay (EMSA): Nuclear protein extracts were obtained from MLE15 mouse lung cells [kind gift from Jeffrey Whitsett, University of Cincinnati] grown in ATCC-formulated RPMI-1640 medium 30-2001 supplemented with 10% fetal bovine serum. Extracts were prepared (NE-PER Nuclear and Cytoplasmic Extraction Reagents, Pierce Biotechnology, Rockford, IL), and assessed by EMSA. The binding reaction mixture containing 3.5 μg of nuclear extract, 2.25 ng of biotin-labeled oligonucleotide probe(s) (C-polymorphic), and 1x binding buffer [in mM: 12 HEPES, pH 7.9, 4 Tris·HCl, pH 7.9, 25 KCl, 50 $MgCl_2$, 1 EDTA, 4 dithiothreitol, 0.2 phenylmethylsulfonyl fluoride,

with 100 ng/μl poly(dI-dC) and 0.05% IGEPAL CA-630] was incubated at 15°C for 30 min in a final volume of 10 μl. Double-stranded oligonucleotides were prepared by annealing complementary synthetic oligonucleotides corresponding to the *Acvr1* promoter region. The nucleotide sequence of the sense strands were:

G-polymorphic: forward GAGAGGCGCTCC**G**GGAGGAGTCTTC

A-polymorphic: forward GAGAGGCGCTCC**A**GGAGGAGTCTTC.

For competition assays, 25-, 50-, or 100-fold excess unlabeled double-stranded variant oligonucleotides were incubated with the extracts (30 min, 23°C) before probe addition. Bound complexes were separated on 6% polyacrylamide gels and blotted onto membrane. The blot was processed with a streptavidin-horseradish peroxidase conjugate-based detection method (LightShift Chemiluminescent Kit, Pierce Biotechnology). Chemiluminescence signal was detected by exposing the blot to film.

Statistics: The commonly used pairwise LD measure, r^2, which is the squared correlation coefficient of alleles at two loci (S3) was calculated for pairwise SNP with a window size of 500 kb across the genome. The associations of the acute lung injury survival time with SNP were tested by using two-sample Student's *t* statistic. A two-sided P value for each SNP was obtained for testing hypothesis of no association between the SNP and survival time and is presented as the negative base-10 logarithmic P value, i.e. -log(P). In addition, a sliding window of three SNP was used to infer haplotypes at each locus. Haplotype-based association analysis was also performed using Student's *t* statistic. In the haplotype analysis, a common variant was first identified and treated as one category and the other rare variants as another category. All analyses were implemented using the R statistical package. One statistical association could occur by chance in a genome-wide LD scan when the threshold is decreased to 4.0 and less than one when the threshold is decreased to 6.0. The threshold for suggestive and significant associations in our whole-genome association analysis can be determined to be 4.0 and 6.0 for -log (P), respectively. All other comparisons of treated mice as compared to strain-matched control or compared to the opposing strain were assessed by Analysis of Variance (ANOVA) followed by an all pairwise multiple comparison procedure (Holm-Sidak method).

Additional Abbreviations (i.e. those not spelled out in the text): **AACS**: acetoacetyl-CoA synthetase, **ABCG1**: ATP-binding cassette, sub-family G (WHITE), member 1, **ACP6**: acid phosphatase 6, lysophosphatidic, **ACVR1:** activin A receptor, type I **ACVR1C**: activin A receptor, type IC, **AKT**: thymoma viral proto-oncogene 1, **ALDH1A7**: aldehyde dehydrogenase family 1, subfamily A7, **ARHGAP15**: Rho GTPase activating protein 15, **BIRC3**: baculoviral IAP repeat-containing 3, **BMP1**: bone morphogenetic protein 1, **BMPR2**: bone morphogenetic protein receptor, type II (serine/threonine kinase), **BNIP3**: BCL2/adenovirus E1B interacting protein 3, **BTG1**: B-cell translocation gene 1, anti-proliferative, **CACNB4**: calcium channel, voltage-dependent, beta 4 subunit, **CACYBP**: calcyclin binding protein, **CCDC48**: coiled-coil domain containing 148 **CDH5**: cadherin 5, **CFLAR**: CASP8 and

FADD-like apoptosis regulator, **CLDN4**: claudin 4, **CREB1**: cAMP responsive element binding protein 1, **CSNK2A1**: casein kinase 2, alpha 1 polypeptide, **CYP1A1**: cytochrome P450, family 1, subfamily a, polypeptide 1, **CYP4F15**: cytochrome P450, family 4, subfamily f, polypeptide 15, **CYR61**: cysteine rich protein 61, **DEDD2**: death effector domain-containing DNA binding protein 2, **DIDO1**: death inducer-obliterator 1, **DNAJA1**: DnaJ (Hsp40) homolog, subfamily A, member 1, **F3**: coagulation factor III, FOS: FBJ osteosarcoma oncogene, **FANCL**: Fanconi anemia, complementation group L, **FAS**: Fas (TNF receptor superfamily member 6), **FKBP4**: FK506 binding protein 4, **FKBP5**: FK506 binding protein 5, **FPR1**: formyl peptide receptor 1, **GALNT13**: UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 13, **GALNT5**: UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 5, **GCLC**: glutamate-cysteine ligase, catalytic subunit, **GPD2**: glycerol phosphate dehydrogenase 2, mitochondrial, **GPR65**: G-protein coupled receptor 65, **GPX2**: glutathione peroxidase 2, **GTF2H2**: general transcription factor II H, polypeptide 2, **HMOX1**: heme oxygenase (decycling) 1, **HSD3B7**: hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 7, **HSP1A1**: heat shock protein 1A, **IL10**: interleukin 10, **IL6**: interleukin 6, **KCNJ3**: potassium inwardly-rectifying channel, subfamily J, member 3, **KITL**: kit ligand, **MAD**: mothers against decapentaplegic, **MAFF**: v-maf musculoaponeurotic fibrosarcoma oncogene family, protein F (avian), **MAP3K1**: mitogen-activated protein kinase kinase kinase 1, **MAP3K7IP2**: mitogen-activated protein kinase kinase kinase 7 interacting protein 2, **MAPK11**: mitogen-activated protein kinase 11, **MAPK12**: mitogen-activated protein kinase 12, **MAPK14**: mitogen-activated protein kinase 14, **MGAT4A:** mannoside acetylglucosaminyltransferase 4, isoenzyme A, **MMP9**: matrix metalloproteinase 9, **MT2**: metallothionein 2, **mTOR**: mechanistic target of rapamycin (serine/threonine kinase), **MVD**: mevalonate (diphospho) decarboxylase**,** **MYCN**: v-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian)**,** **NFAT5**: nuclear factor of activated T-cells 5, **NFE2L2**: nuclear factor, erythroid derived 2, like 2, **NQO1**: NAD(P)H dehydrogenase, quinone 1, **NR2F2**: nuclear receptor subfamily 2, group F, member 2, **NR4A2**: nuclear receptor subfamily 4, group A, member 2, **p53 (TRP53)**: transformation related protein 53 **PERP**: PERP, TP53 apoptosis effector, **PRKAR2B**: protein kinase, cAMP dependent regulatory, type II beta, **PRKCZ**: protein kinase C, zeta, **PRMT2**: protein arginine N-methyltransferase 2, **PSCDBP (CYTIP)** cytohesin 1 interacting protein, **RAC1**: RAS-related C3 botulinum substrate 1, **RAR**: retinoic acid receptor, alpha, **RARB**: retinoic acid receptor, beta, **RBMS1**: RNA binding motif, single stranded interacting protein 1, **RBP1**: retinol binding protein 1, cellular, **RFWD2**: ring finger and WD repeat domain 2, **RPAP2**: RNA polymerase II associated protein 2, **SMAD:** MAD homolog 1 (Drosophila), **SMARCB1**: SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1, **SPP1**: secreted phosphoprotein 1, **STAM2**: signal transducing adaptor molecule (SH3 domain and ITAM motif) 2, **STIP1**: stress-induced phosphoprotein 1, **STK17B**: serine/threonine kinase

17b (apoptosis-inducing)**, TANC1**: tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 1, **TGFB1**: transforming growth factor, beta 1, **TGFBR3**: transforming growth factor, beta receptor III, **TGIF1**: TGFB-induced factor homeobox 1, **THBS1**: thrombospondin 1, **TNN**: tenascin N, **TXNRD1**: thioredoxin reductase 1.

Literature Cited

E1. Smyth GK. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;3, Article3.

E2. Sartor MA, Tomlinson CR, Wesselkamper SC, Sivaganesan S, Leikauf GD, Medvedovic M. Intensity-based hierarchical Bayes method improves testing for differentially expressed genes in microarray experiments, *BMC Bioinform* 2006; 7: 538.

E3. Hedrick PW. Gametic disequilibrium measures: proceed with caution. *Genetics* 1987;117: 331-341.

Supplemental Figures

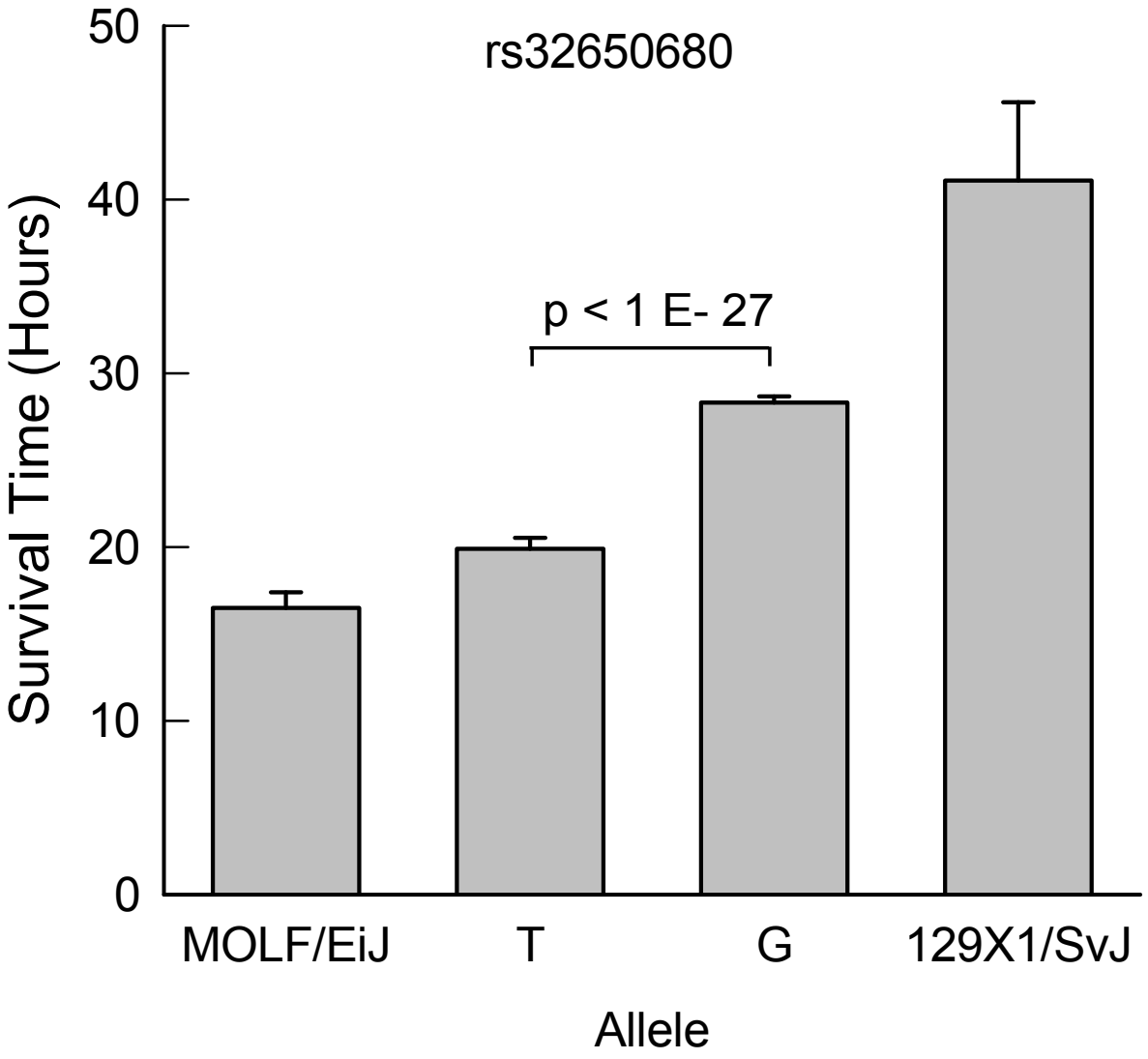

Figure E1. Assessment of the phenotypic difference in mean survival time between the most sensitive (MOLF/EiJ) and resistant (129X1/SvJ) mouse strains resulting from a single nucleotide polymorphism (SNP) (rs32650680) association in ring finger WD repeat domain 2. The mean survival time was determined for mice that have either the T (mean = 19.9 ± 0.4 h, n = 137 mice) or G (mean = 28.3 ± 0.7 h, n = 228 mice) allele. The difference between these groups was then compared to the difference of the means of the most sensitive and resistant of the 40 mouse strains exposed to 10 ppm acrolein. In this case, the difference explained was 34%. The -log (P) value = 5.71 for this SNP. Value are means ± standard errors and the significance of the difference (p value) between the T and G allele means was determined by analysis of variance (ANOVA) with an all pairwise multiple comparison procedure (Holm-Sidak method).

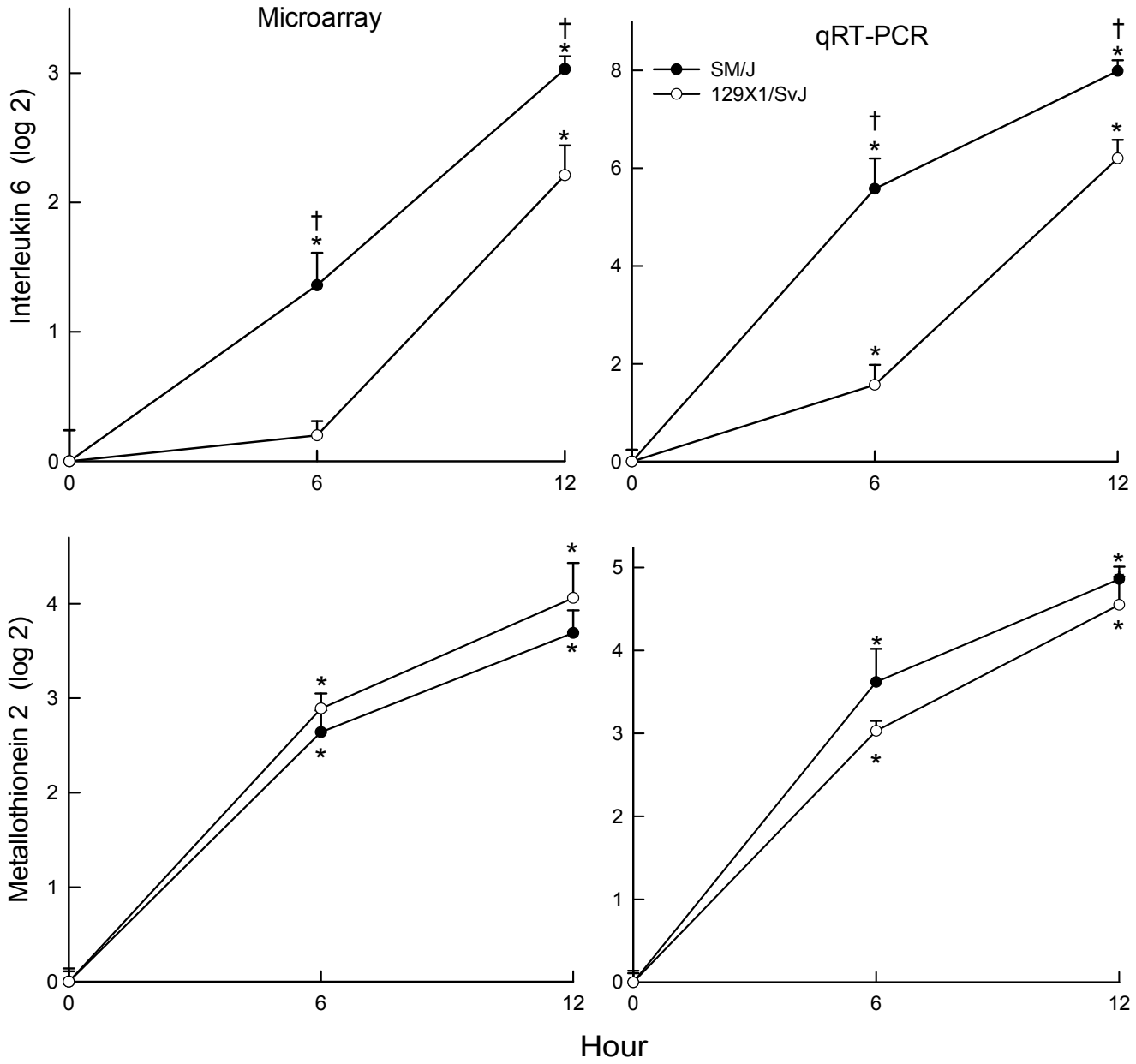

Figure E2. Transcriptional changes of interleukin 6 and metallothionein 2 during acute lung injury. Mice were exposed to 10 ppm acrolein for 0 (filtered air control), 6, or 12 hours, anesthetized, and lung tissue frozen in liquid nitrogen. Lung mRNA was isolated and (**Top**) interleukin 6 or (**Bottom**) metallothionein 2 transcript levels were determined by (**Right**) microarray or (**Left**) quantitative real time polymerase chain reaction. Values are mean ± SE (n=7-8 mice). *****Significantly different from strain-matched control as determined by analysis of variance (ANOVA) with an all pairwise multiple comparison procedure (Holm-Sidak method). †Significantly different between the sensitive SM/J and resistant 129X1/SvJ mouse strains as determined by analysis of variance (ANOVA) with an all pairwise multiple comparison procedure (Holm-Sidak method).

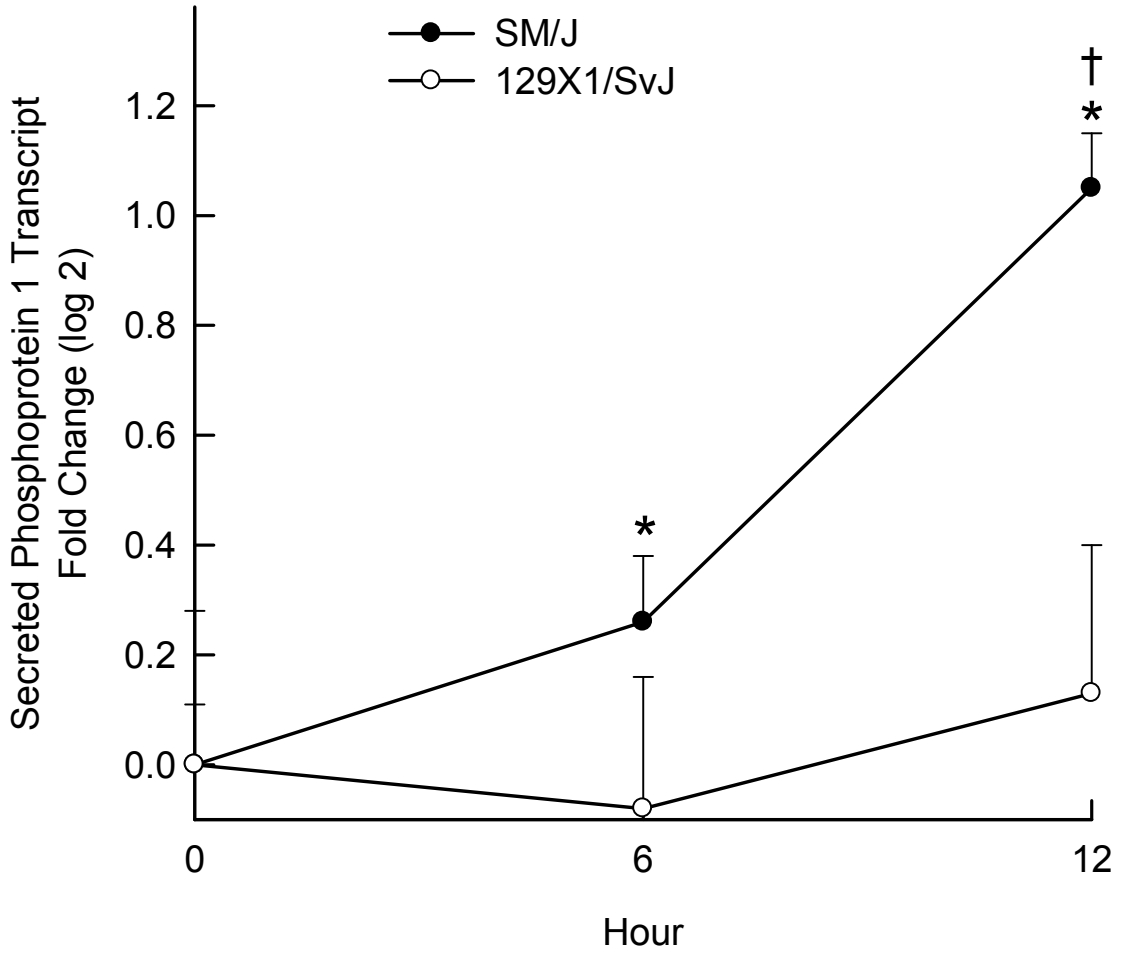

Figure E3. Transcript levels of secreted phosphoprotein 1 (SPP1), a transforming growth factor beta target during acute lung injury in mice. Mice were exposed to 10 ppm acrolein for 0 (filtered air control), 6, or 12 hours, anesthetized, and lung tissue frozen in liquid nitrogen. Lung mRNA was isolated and transcript levels determined by quantitative real time polymerase chain reaction. Values are mean ± SE (n=7-8 mice). *****Significantly different from strain-matched control as determined by analysis of variance (ANOVA) with an all pairwise multiple comparison procedure (Holm-Sidak method). †Significantly different between the sensitive SM/J and resistant 129X1/SvJ mouse strains as determined by analysis of variance (ANOVA) with an all pairwise multiple comparison procedure (Holm-Sidak method).

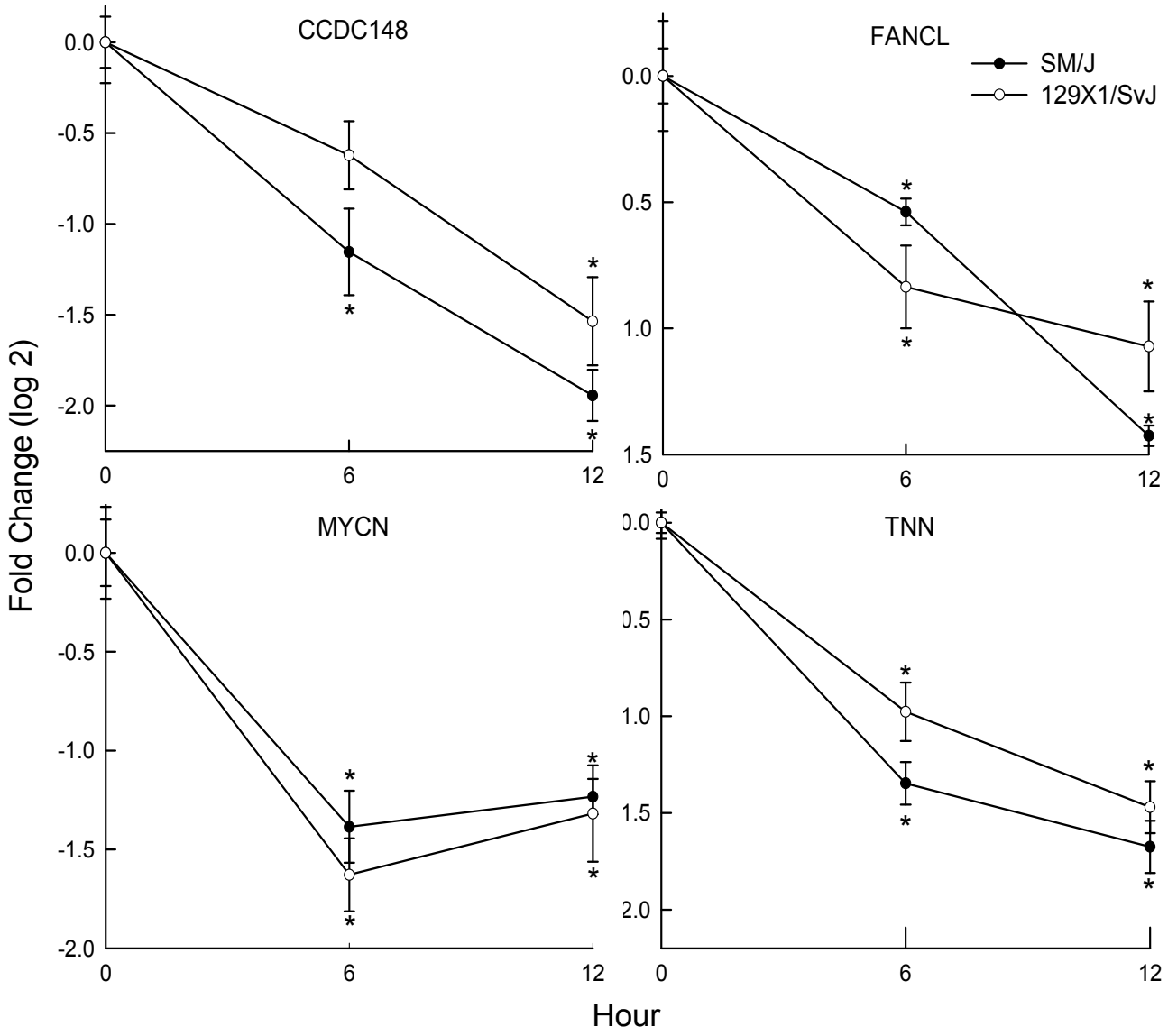

Figure E4. Transcript levels of 6 candidate genes that differed between the SM/J and 129X1/SvJ mouse strains following acrolein exposure. Mice were exposed to filtered air (Control, 0h), or to acrolein (10 ppm) for 6 or 12 h, lung mRNA isolated, and transcript expression levels determined by quantitative real time polymerase chain reaction. ***Abbreviations:*** **ACVR1**: activin A receptor, type 1, **ARHGAP15**: Rho GTPase activating protein 15, **CACNB4**: calcium channel, voltage-dependent, beta 4 subunit, **CACYBP**: calcyclin binding protein, **RFWD2**: Ring finger and WD repeat domain 2, **TGFBR3**: transforming growth factor, beta receptor III. Values are mean ± SE of the transcript level of SM/J (n=8) as compared to the 129X1/SvJ (n=7-8). *****Significantly different from strain-matched control as determined by analysis of variance (ANOVA) with an all pairwise multiple comparison procedure (Holm-Sidak method).

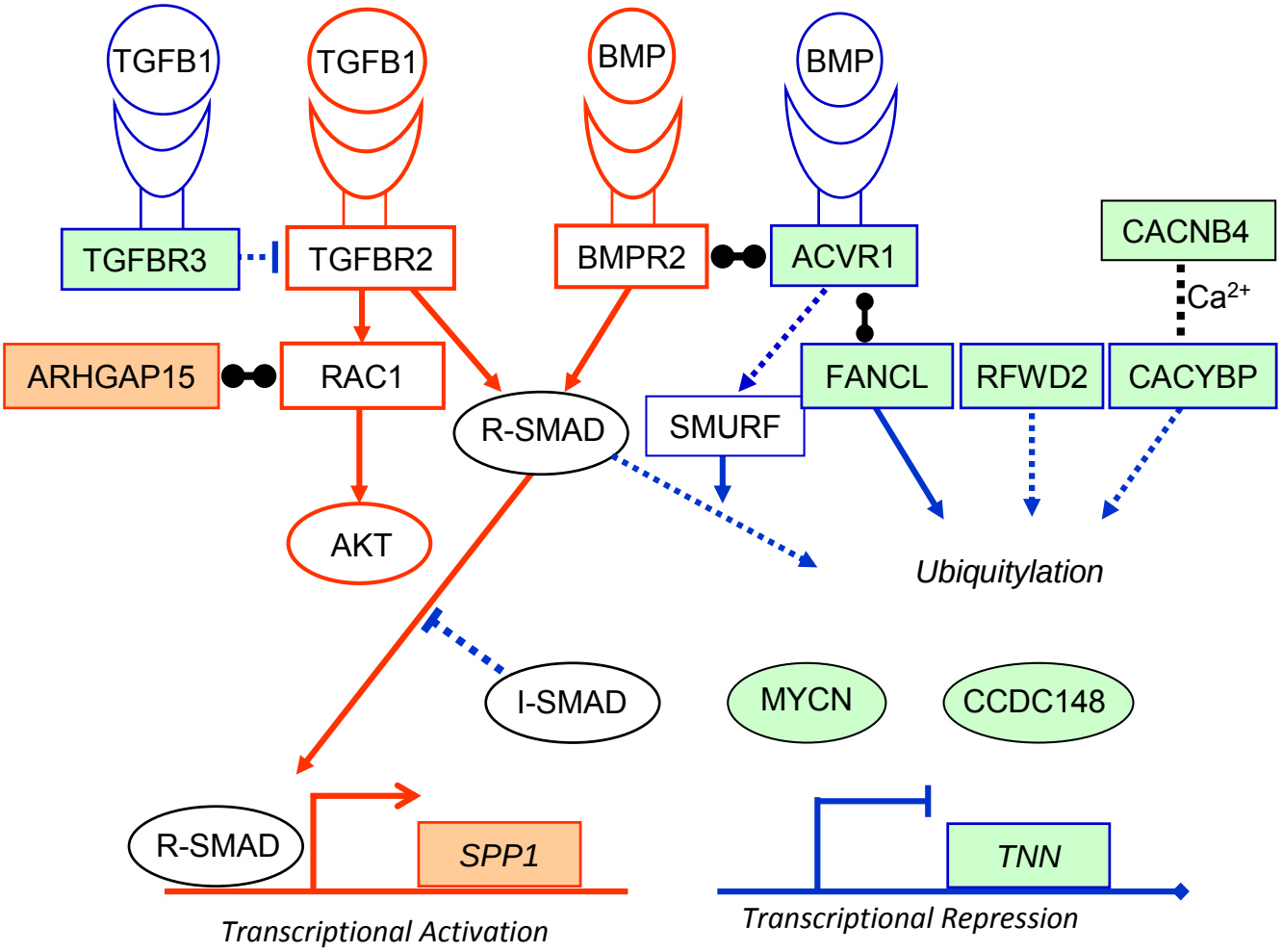

Figure E5. Interactions among candidate genes associated with transforming growth factor, beta (TGFB)/bone morphogenetic protein (BMP) signaling associated with acrolein-induced acute lung injury. Two of the candidates (TGFBR3 and ACVR1) are receptors for TGFB and BMP ligands. In turn, TGFBR2 interacts with RAC1, which can be regulated by ARHGAP15, and/or stimulates regulatory R-SMAD transcription factors that can either activate or inhibit transcription of target genes. ACVR1 partners with BMPR2 to interact with FANCL or SMURF proteins that modulate ubiquitination. Additional proteins with ubiquitin-protein ligase activity include RFWD2 and CACYBY. Transcripts that are colored **Salmon** increased or **Green** decreased during acrolein exposure. Pathways that are lined in **Red** activate or **Blue** inhibit TGF/BMP signaling. ***Abbreviations:*** **ACVR1**: activin A receptor, type 1, **AKT**: thymoma viral proto-oncogene 1, **ARHGAP15**: Rho GTPase activating protein 15, **BMP**: bone morphogenetic protein, **BMPR2**: bone morphogenetic protein receptor, type II, **CACNB4**: calcium channel, voltage-dependent, beta 4 subunit, **CACYBP**: calcyclin binding protein, **CCDC148**: coiled-coil domain containing 148, **FANCL**: Fanconi anemia, complementation group L, **MYCN**: v-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian), **RAC1**: RAS-related C3 botulinum substrate 1, **RFWD2**: Ring finger and WD repeat domain 2, **I-SMAD**: inhibitory mothers against decapentaplegic homologs (SMADs), **R-SMAD**: regulator SMADs, **SMURF**: SMAD specific E3 ubiquitin protein ligases, **SPP1**: secreted phosphoprotein 1, **TGFB1**: transforming growth factor, beta 1, **TGFBR2**: transforming growth factor, beta receptor II, **TGFBR3**: transforming growth factor, beta receptor III, **TNN**: tenascin N.

Supplemental Tables

Table E1. Single Nucleotide Polymorphism (SNP) associations within candidate genes

Table E2. Number of unique transcripts altered in lung from sensitive (SM/J) and resistant (129X1/SvJ) mouse strains exposed for 6 or 12 h compared to control (0 h)

Table E3. Single Nucleotide Polymorphism (SNP) genotyping of activin A receptor, type I promoter region

Table E4. Putative interactions of candidate genes in acrolein-induced acute lung injury in mice

TABLE E1. SINGLE NUCLEOTIDE POLYMORPHISM (SNP) ASSOCIATIONS IN CANDIDATE GENES

dbSNP	- log (P)	Chr	Position	Alleles	Allelic Frequency	Phenotype (%)	Candidate Gene
rs31318131	5.78	1	161197664	A/G	0.35	30.2	*Rfwd2*
rs32650680	5.71	1	161199419	T/G	0.38	34.1	*Rfwd2*
rs32651097	5.71	1	161252359	C/G	0.38	31.8	*Rfwd2*
rs27914711	5.27	2	52317165	A/G	0.27	18.8	*Cacnb4*
(liftover)	5.30	2	52417344	C/T	ND	ND	*Cacnb4*
rs3676875	5.34	2	52418654	T/C	0.25	17.3	*Cacnb4*
rs27943678	5.15	2	54304132	T/C	0.43	18.1	*Galnt13*
rs27926230	5.05	2	54309865	C/A	0.29	21.0	*Galnt13*
(liftover)	5.13	2	54642166	T/C	ND	ND	*Galnt13*
rs27941284	5.95	2	58265297	T/G	0.40	27.1	*Acvr1*
rs27924920	5.34	2	58298494	A/G	0.25	18.9	*Acvr1*
rs13476523	6.94	2	58348075	T/C	0.33	19.4	*Acvr1*
rs3667874	6.17	2	58352456	G/A	0.44	24.3	*Acvr1*
rs13476525	7.69	2	58748696	T/C	0.20	18.8	*Ccdc148*
rs29508317	4.35	5	107548115	T/C	0.33	16.8	*Tgfbr3*
rs32049406	5.33	5	108052320	C/T	0.54	31.1	*Rpap2*
rs29731711	5.87	5	108056144	C/T	0.41	23.2	*Rpap2*

Abbreviations: **dbSNP**: Single nucleotide polymorphism database identification number, **- log (P)**: - log of the association probability, **Chr**: chromosome, **Position**: basepair location of SNP, **Alleles**: nucleotide base polymorphic "minor" allele first, **Allelic Frequency**: Frequency of the minor allele in 365 mice from 40 mouse strains, **Phenotype (%)**: Percentage phenotype (mean survival time) difference explained by mice with either allele as a percentage of the difference of 2 most polar strains, **Candidate Gene**: Gene in which SNP associations are located, ***Rfwd2***: Ring finger WD repeat domain 2, ***Cacnb4***: Calcium channel voltage-dependent, beta 4 subunit, ***Galnt13***: UDP-N-acetyl-alpha-D-galactosamine:N-acetylgalactosaminyltransferase 13, ***Acvr1***: Activin A receptor, type 1, ***Ccdc148***: coiled-coil domain containing 148, ***Tgfbr3***: Transforming growth factor, beta receptor III, ***Rpap2***: RNA polymerase II associated protein 2 **ND**: not determined due to a lack of genotyping

TABLE E2. NUMBER OF UNIQUE TRANSCRIPTS ALTERED IN THE LUNG OF SENSITIVE (SM/J) AND RESISTANT (129X1/SVJ) MOUSE STRAINS EXPOSED TO ACROLEIN FOR 6 OR 12H COMPARED TO CONTROL (0H) (n = 5 MICE/TIME)

	SM/J Mean		129X1/SvJ	
	6 vs 0h	12 vs 0h	6 vs 0h	12 vs 0h
$p < 0.05$				
Increased	2414	3165	2042	2779
Decreased	2857	3594	2240	3209
$p < 0.01$ with log 2 $\geq$ 1.0				
Increased	599	1191	608	933
Decreased	312	1050	418	800

TABLE E3. SINGLE NUCLEOTIDE POLYMORPHISM (SNP) GENOTYPING OF ACTIVIN A RECEPTOR, TYPE I PROMOTER REGION

Mbp location (Build 37)	129X1/SvJ	A/J	BALB/cByJ	BALB/cJ	BTBR T+tf/J	C3H/HeJ	C57BL/10J	C57BL/6J	C57BLKS/J	C57BR/cdJ	CBA/J	CZECHII/EiJ	DBA/1J	DBA/2J	FVB/NJ	I/LnJ	KK/HlJ	LG/J	MOLF/EiJ	MRL/MpJ	NOD/ShiLtJ	NON/ShiLtJ	NZW/LacJ	PL/J	RIIIS/J	SJL/J	SWR/J	WSB/EiJ	db SNP rs	Observed
02 58.368612	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	rs33408603	A/G
02 58.368619	G	G	G	G	C	G	G	G	G	G	G	C	G	G	G	G	G	G	C	G	G	G	G	G	G	G	G	C	rs6406068	C/G
02 58.368637	G	G	G	G	A	G	G	G	G	G	G	A	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	A	rs6406107	A/G

TABLE E4. PUTATIVE INTERACTIONS OF CANDIDATE GENES IN ACROLEIN-INDUCED ACUTE LUNG INJURY IN MICE

Protein 1	Protein 2	Interaction	Function	Putative Consequence in SM/J	Reference
TGFB1	TGFBR3	Ligand binding	Association TGFBR3 with TGFBR2	Less activation of TGFBR3	Blobe GC, Schiemann WP, Pepin MC, Beauchemin M, Moustakas A, Lodish HF, O'Connor-McCourt MD. Functional roles for the cytoplasmic domain of the type III transforming growth factor beta receptor in regulating transforming growth factor beta signaling. J Biol Chem 2001 Jul 6;276(27):24627-37.
TGFB1	TGFBR2	Ligand binding	Decreased activation of TGFBR2	More activation of TGFBR2	Moustakas A, Pardali K, Gaal A, Heldin CH. Mechanisms of TGF-beta signaling in regulation of cell growth and differentiation. Immunol Lett 2002 Jun 03;82(1-2):85-91.
BMP	BMPR2	Ligand binding	SMAD signaling	Possibe augmentation of SMAD	Shi Y, Massagué J.Mechanisms of TGF-beta signaling from cell membrane to the nucleus. Cell. 2003 Jun 13;113(6):685-700
BMP	ACVR1	Ligand binding	SMAD Signaling	Possibe dysregulation of SMAD	Song GA, Kim HJ, Woo KM, Baek JH, Kim GS, Choi JY, Ryoo HM.Molecular consequences of the ACVR1(R206H) mutation of fibrodysplasia ossificans progressiva. J Biol Chem. 2010 Jul 16;285(29):22542-53.
TGFBR3	TGFBR2	Association of receptors	Decrease in TGF-beta signaling	Less inhibition of TGFBR2	Eickelberg O, Centrella M, Reiss M, Kashgarian M, Wells RG. Betaglycan inhibits TGF-beta signaling by preventing type I-type II receptor complex formation. Glycosaminoglycan modifications alter betaglycan function. J Biol Chem 2002 Jan 4;277(1):823-9.
TGFBR2	RAC1	Rho GTPase signaling	Cytoskeleton rearrangement, cell proliferation	More activatiom of RAC1	Varon C, Rottiers P, Ezan J, Reuzeau E, Basoni C, Kramer I, Génot E. TGFbeta1 regulates endothelial cell spreading and hypertrophy through a Rac-p38-mediated pathway. Biol Cell. 2008 Sep;100(9):537-50
TGFBR2	SMAD	Phosphorylation of SMAD2/3	SMAD signaling	More activation of SMAD signaling	Nakao A, Imamura T, Souchelnytskyi S, Kawabata M, Ishisaki A, Oeda E, Tamaki K, Hanai J, Heldin CH, Miyazono K, ten Dijke P. TGF-beta receptor-mediated signalling through Smad2, Smad3 and Smad4. EMBO J 1997 Sep 1;16(17):5353-62.
ACVR1	BMPR2	Protein-Protein interaction	BMP signaling	Possibe dysregulation of SMAD	Derynck R, Zhang YE. Smad-dependent and Smad-independent pathways in TGF-beta family signalling. Nature. 2003 Oct 9;425(6958):577-84.
ACVR1	FANCL	ACVR1 binds FANCL	Decreased FANCD ubiquitylation	Less DNA repair	Barrios-Rodiles M, Brown KR, Ozdamar B, Bose R, Liu Z, Donovan RS, Shinjo F, Liu Y, Dembowy J, Taylor IW, Luga V, Przulj N, Robinson M, Suzuki H, Hayashizaki Y, Jurisica I, Wrana JL. High-throughput mapping of a dynamic signaling network in mammalian cells. Science 2005 03 11;307(5715):1621-5.
CACNB4	CACYBP	Decreases calcium entry	Decrease CACYBP mediated ubiquitylaton	Inappropriate ubiquitylation of damaged proteins	Matsuzawa SI, Reed JC. Siah-1, SIP, and Ebi collaborate in a novel pathway for beta-catenin degradation linked to p53 responses. Mol Cell 2001 May 1;7(5):915-26.
ARHGAP15	RAC1	Protein-Protein interaction	Cytoskeleton rearrangement		Seoh ML, Ng CH, Yong J, Lim L, Leung T. ArhGAP15, a novel human RacGAP protein with GTPase binding property. FEBS Lett 2003 03 27;539(1-3):131-7.
RAC1	AKT	Phosphorylation of AKT	Decreased AKT1 activation	Disruption of cytoskeleton and barrier function, and augmented cell death	Singleton PA, Chatchavalvanich S, Fu P, Xing J, Birukova AA, Fortune JA, Klibanov AM, Garcia JG, Birukov KG. Akt-mediated transactivation of the S1P1 receptor in caveolin-enriched microdomains regulates endothelial barrier enhancement by oxidized phospholipids.Circ Res. 2009;104:978-986.
SMURF	SMAD	Protein-Protein interaction	Decreased SMAD ubiquiylation	Augmented SMAD signaling	Zhang Y, Chang C, Gehling DJ, Hemmati-Brivanlou A, Derynck R. Regulation of Smad degradation and activity by Smurf2, an E3 ubiquitin ligase. Proc Natl Acad Sci U S A 2001 Jan 30;98(3):974-9.
FANCL	Ubiquitin Protein Complex	Protein-Protein interaction	Decreased FANCD ubiquitylation	Increased susceptibility to DNA damage	Jacquemont C, Taniguchi T. The Fanconi anemia pathway and ubiquitin. BMC Biochem 2007 01 1;8 Suppl 1:S10. and Cole AR, Lewis LP, Walden H.The structure of the catalytic subunit FANCL of the Fanconi anemia core complex. Nat Struct Mol Biol. 2010 Mar;17(3):294-8.
RFWD2	JUN	Protein-Protein interaction	Decreased RFWD2	Increased JUN and FOXO1, and repressed SFTPB	Wertz IE, O'Rourke KM, Zhang Z, Dornan D, Arnott D, Deshaies RJ, Dixit VM. Human De-etiolated-1 regulates c-Jun by assembling a CUL4A ubiquitin ligase.Science. 2004 Feb 27;303(5662):1371-4
CACYBP	S100A6	Protein-Protein interaction	Decreased calcium mediated ubiquitylation	Inappropriate ubiquitylation of damaged proteins	Filipek A, Jastrzebska B, Nowotny M, Kuznicki J. CacyBP/SIP, a calcyclin and Siah-1-interacting protein, binds EF-hand proteins of the S100 family.J Biol Chem. 2002 Aug 9;277(32):28848-52
SMAD	SPP1	Transcriptoinal activation	Activation of natural killer T Cells	Augmentation of cell death	Zhang ZX, Shek K, Wang S, Huang X, Lau A, Yin Z, Sun H, Liu W, Garcia B, Rittling S, Jevnikar AM. Osteopontin expressed in tubular epithelial cells regulates NK cell-mediated kidney ischemia reperfusion injury. J Immunol. 2010 Jul 15;185(2):967-73.
MYCN	TNN	Transcriptoinal repression	Decreased TNN	Reduce cell-matrix adhesion	Scherberich A, Tucker RP, Degen M, Brown-Luedi M, Andres AC, Chiquet-Ehrismann R. Tenascin-W is found in malignant mammary tumors, promotes alpha8 integrin-dependent motility and requires p38(MAPK) activity for BMP-2 and TNF-alpha induced expression in vitro. Oncogene 2005;24:1525-32.
CCDC148	Unknown	Unknown	Unknown	Vascular dysregulation	Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007 Jun 7;447(7145):661-78.