



Editing out five *Serpina1* paralogs to create a mouse model of genetic emphysema

Florie Borel^a, Huaming Sun^a, Marina Zieger^a, Andrew Cox^a, Brynn Cardozo^a, Weiyang Li^a, Gabriella Oliveira^a, Airiel Davis^b, Alisha Gruntman^{a,b}, Terence R. Flotte^{a,c}, Michael H. Brodsky^d, Andrew M. Hoffman^b, Mai K. Elmallah^{a,c}, and Christian Mueller^{a,c,1}

^aHorae Gene Therapy Center, University of Massachusetts Medical School, Worcester, MA 01605; ^bCummings School of Veterinary Medicine, Tufts University, Grafton, MA 01536; ^cDepartment of Pediatrics, University of Massachusetts Medical School, Worcester, MA 01605; and ^dDepartment of Molecular, Cell, and Cancer Biology, University of Massachusetts Medical School, Worcester, MA 01605

Edited by Michael J. Welsh, Howard Hughes Medical Institute, Iowa City, IA, and approved January 22, 2018 (received for review August 2, 2017)

Chronic obstructive pulmonary disease affects 10% of the worldwide population, and the leading genetic cause is α -1 antitrypsin (AAT) deficiency. Due to the complexity of the murine locus, which includes up to six *Serpina1* paralogs, no genetic animal model of the disease has been successfully generated until now. Here we create a quintuple *Serpina1a-e* knockout using CRISPR/Cas9-mediated genome editing. The phenotype recapitulates the human disease phenotype, i.e., absence of hepatic and circulating AAT translates functionally to a reduced capacity to inhibit neutrophil elastase. With age, *Serpina1* null mice develop emphysema spontaneously, which can be induced in younger mice by a lipopolysaccharide challenge. This mouse models not only AAT deficiency but also emphysema and is a relevant genetic model and not one based on developmental impairment of alveolarization or elastase administration. We anticipate that this unique model will be highly relevant not only to the preclinical development of therapeutics for AAT deficiency, but also to emphysema and smoking research.

Serpina1 | alpha-1 antitrypsin | emphysema | CRISPR | mouse model

Chronic obstructive pulmonary disease (COPD) is the third leading cause of mortality worldwide (1), and it affects about 10% of the world population (2). While cigarette smoking and exposure to air particulates are well-characterized environmental risk factors, α -1 antitrypsin (AAT) deficiency is the most common known genetic factor (3). AAT deficiency is characterized by mutations in the SERPINE Proteinase Inhibitor family A member 1 (*SERPINA1*) gene. These mutations result in reduced or inexistent levels of serum AAT protein (4), a protease inhibitor whose physiological target is neutrophil elastase (NE). Low or no serum AAT therefore results in the unopposed action of neutrophil elastase on the alveolar interstitium thereby leading to destruction of the alveolar walls and lung parenchyma. Clinically, this results in pulmonary emphysema, which is characterized by fewer, larger air sacs with less recoil and decreased gas exchange, leading to increasing breathlessness.

The current standard of care is weekly or biweekly intravenous infusion of recombinant AAT protein. Unfortunately, because this disease is underdiagnosed, many patients start treatment after their respiratory function is significantly and irreversibly compromised. In addition, lifelong weekly or biweekly infusions remain a burden for patients, and alternative therapeutic options allowing for less frequent interventions are being actively investigated, including gene therapy (5–8) and genome editing (9). However, the development of new therapeutics is hindered by the absence of an animal model (10). Current study goals for protein replacement therapies are based on the FDA-defined endpoint of 11 μ M serum AAT and not on a physiologically relevant endpoint such as respiratory measurements, which may limit approval of functional candidates. The absence of a mouse model is due to the presence of *Serpina1* paralogs in the murine genome (11), which made the generation of a full knockout a very difficult target using traditional methods. In addition, some reports have dissuaded these efforts due to the possibility of an embryonic role for these genes in

mice. In fact a murine *Serpina1a* knockout was recently described to be embryonically lethal (12), as well as a murine *Serpina1b* knockout (13). However, it is unclear why mice not only amplified the gene but also developed an embryonic role for it and why deleting only one of several highly similar murine paralogs would be lethal while *SERPINA1* null humans are viable. Prompted by this incongruity and the advent of CRISPR technology, we sought to confirm these surprising findings by generating a full knockout using guide RNAs (gRNAs) that target a sequence in exon 2 that is conserved in all *Serpina1* genes. Here, we demonstrate the successful generation of a murine complete *Serpina1a-e* knockout with biallelic editing of five genes in a single round of zygote injections. This mouse model has a normal lifespan but presents with a respiratory phenotype that recapitulates many aspects of the human disease and therefore constitutes a true and robust model in which to validate novel therapeutics and investigate the biology of emphysema.

Results

Generation of the *Serpina1a-e* Knockout. Due to gene amplification events, six *Serpina1* genes exist in the murine genome. These are designated in the current nomenclature as *Serpina1a* through *Serpina1f*, although only *Serpina1a-e* are expressed in the liver. The genes are 11 kb long and span about 230 kb on chromosome 12. In our mouse strain of choice, C57BL/6J, those five *Serpina1* genes (*a-e*) are present and expressed (11). The corresponding proteins are designated as α -1 antitrypsin (AAT) 1-1 through 1-5. Of these, AAT 1-1 and AAT 1-2 inhibit both neutrophil and

Significance

Chronic obstructive pulmonary disease affects 10% of the worldwide population, and the leading genetic cause is a genetic disease, α -1 antitrypsin (AAT) deficiency. Humans have only one gene that codes for the AAT protein, but mice have up to six, which made it impossible for decades to create a mouse model of the disease. Here we succeeded in creating this mouse model using CRISPR technology to target all of the mouse genes at once. Importantly, this mouse model spontaneously develops lung disease and recapitulates many aspects of the human disease. We anticipate that this model will be highly relevant not only to the preclinical development of therapeutics for AAT deficiency, but also to emphysema and smoking research.

Author contributions: F.B., M.H.B., and C.M. designed research; F.B., M.Z., A.C., B.C., W.L., G.O., A.D., A.G., M.H.B., and C.M. performed research; F.B., H.S., M.Z., A.C., B.C., W.L., G.O., A.D., T.R.F., A.M.H., M.K.E., and C.M. analyzed data; and F.B. and C.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹To whom correspondence should be addressed. Email: chris.mueller@umassmed.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1713689115/-DCSupplemental.

Published online February 16, 2018.

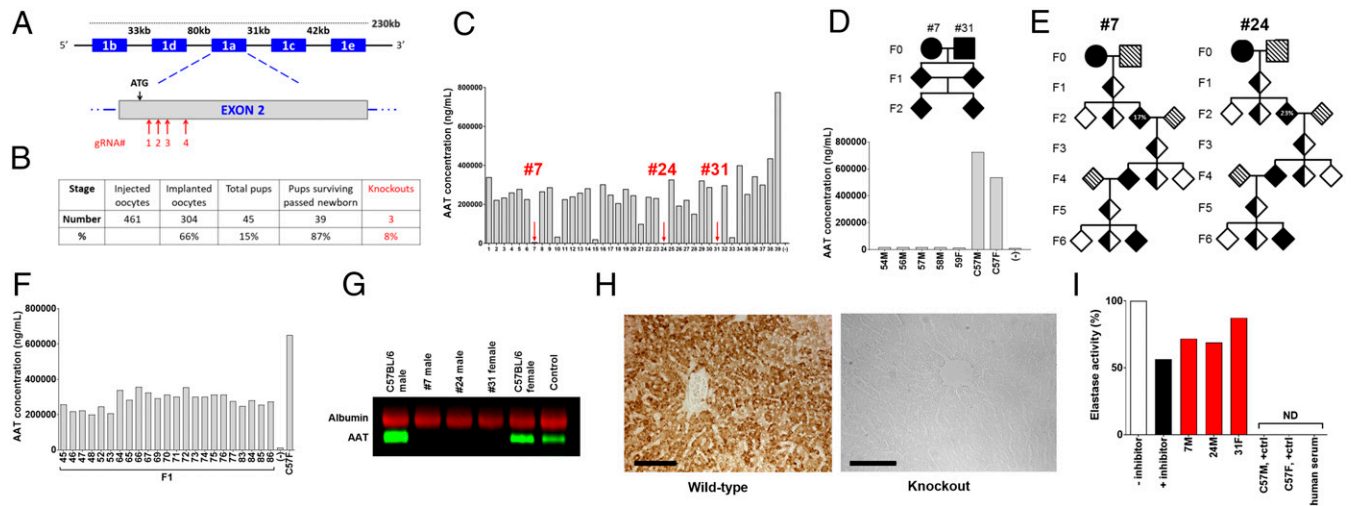


Fig. 1. Generation of the *Serpina1* knockout. (A) The five *Serpina1* genes span a 230-kb region on chromosome 12 and their 11-kb sequence is highly homologous. Guide RNAs (gRNAs) were designed to target a conserved region in exon 2. (B) Statistics on the knockout generation from zygote injection to identification of the knockout line founders. (C) Pups were screened by murine AAT ELISA, and the three founders #7, #24, and #31 were identified by undetectable protein levels, indicated by red arrows. –, negative control. (D) Breeding male founder #7 with female founder #31 resulting in 100% of the progeny being knockouts, as determined by serum AAT ELISA, demonstrated germline transmission. (E) Three independent founder backcross lines were generated by backcrossing: line 7A from founder #7, line 24B from founder #24, and line 31C from founder #31 (line 31C is currently at F4 only). (F) Backcrossing male founders #7 and #24 resulting in 100% of the progeny being heterozygous, as determined by serum AAT ELISA, demonstrated genetic linkage (pups 45–67, line 7A; pups 69–86, line 24B; –, negative control; C57F, wild-type mouse, positive control). The serum of the three founders is devoid from AAT protein, as confirmed by Western blot (G). (H) Immunohistochemistry shows absence of AAT in liver tissue from knockout animals. These representative images show hepatocytes surrounding a blood vessel. (Scale bar, 100 μ m.) (I) Serum from the three founders has reduced antielastase activity, as determined by quantification of elastase activity. The – inhibitor bar is an internal assay control representing an uninhibited reaction, and its value is normalized at 100. C57M, male wild-type mouse; 7M, male founder #7; 24M, male founder #24; 31F, female founder #31; C57F, female wild-type mouse; and AAT control, commercial pooled mouse serum.

pancreatic elastases, while AAT 1-3 and AAT 1-4 do not, and AAT 1-5 is not well characterized (14). Four gRNAs were designed in a region conserved in all *Serpina1* paralogs, exon 2 (Fig. 1A), and validated in vitro with a GFP-based single strand annealing assay. Genetically modified mice were generated using previously described methods (15). Briefly, zygotes were microinjected with Cas9 mRNA and all four gRNAs, and subsequently implanted in 13 pseudopregnant females. A total of 45 pups were born, of which 87% survived weaning (Fig. 1B) and their sera was screened for total murine AAT load by ELISA using a polyclonal antibody. Out of the 39 viable animals, 3 were identified as complete biallelic *Serpina1a-e* knockouts due to no detectable expression of AAT (Fig. 1C), two males and one female (subsequently referred to as #7, #24, and #31). Germline transmission was demonstrated when crossing founders #7 and #31 led 100% of the progeny to have undetectable levels of circulating AAT (Fig. 1D). To get rid of hypothetical undesired genomic alterations due to CRISPR-related off-target effects, founders #7 and #24 were independently backcrossed three times (Fig. 1E) and founder #31 has currently been backcrossed twice. This led to the generation of three independent lines of *Serpina1a-e* knockouts subsequently referred to as 7A, 24B, and 31C for founders #7, #24, and #31, respectively. Genetic linkage of the disrupted alleles was demonstrated when the first backcross of founders #7 and #24 led to a 100% heterozygous progeny (Fig. 1F), as determined by AAT serum levels. *Serpina1a-e* knockout mice have a normal lifespan, ability to breed, and gender distribution.

Phenotypic and Functional Characterization. AAT is produced in the liver, secreted in the bloodstream, and eventually reaches the lungs, which is its main site of action. In the lungs, AAT protects the lung alveoli and interstitial elastin from destruction by inactivating neutrophil elastase. To assess whether the protein was present in the circulation, blood serum was sampled and analyzed by Western blot using an antimouse AAT primary antibody. The three founders demonstrated undetectable levels of AAT (Fig. 1G), confirming the results of the initial ELISA screening (Fig. 1C). Next, liver sections

were subjected to immunohistochemical staining. While a strong staining reflects the high abundance of hepatic AAT in wild-type mice, the *Serpina1a-e* knockout mice revealed a complete lack of staining, denoting absence of the protein (Fig. 1H). Finally, we assessed the biochemical consequences of the lack of circulating AAT by analyzing the ability of the sera to neutralize elastase activity. Blood serum was incubated with elastase and an analog chemical substrate, and relative elastase activity was derived from a colorimetric quantification of the kinetics of substrate degradation. While serum from wild-type male and female C57BL/6 mice exhibited a high capacity to inhibit elastase, sera from the three founders showed a reduced capacity to do so (Fig. 1I, red bars). The reduced capacity of the serum samples from the three founders to inhibit elastase was similar to the results obtained when using a commercial elastase inhibitor (Fig. 1I, black bar). This indicates that absence of serum AAT shown in Fig. 1C and G translates to a loss of its physiological antielastase function. Any remaining antielastase activity in the sera of the founders might be due to other serum serine proteinase inhibitors (serpins) with antielastase activity.

Genomic and Transcriptomic Characterization. Genomic editing alterations were characterized by targeted sequencing. Briefly, high molecular weight genomic DNA (gDNA) was isolated from liver tissue of one mouse per founder line. This gDNA was subsequently sheared to a size of \sim 6 kb, and fragments of interest were captured using probes specific to the 230-kb region of interest (Fig. 1A). The \sim 2-kb captured fragments were sequenced by long-read PacBio sequencing, to allow for discrimination between highly homologous *Serpina1* paralogs. Analysis revealed excisions of up to 90 kb in lines 7A and 24B, and up to 76 kb in line 31C (Fig. 2 and Figs. S1–S3). In line 24B, complementary targeted locus amplification (TLA) sequencing revealed a small excision of 73 bp in *Serpina1b*, at the location of gRNA#2 (Figs. S5 and S6).

To confirm the excisions and nonsense-mediated decay (NMD)-causing frameshifts, RNA was isolated from liver tissue of two of the three founder lines for a transcriptome analysis. Four mice from the F4 generation of line 7A and four mice from the

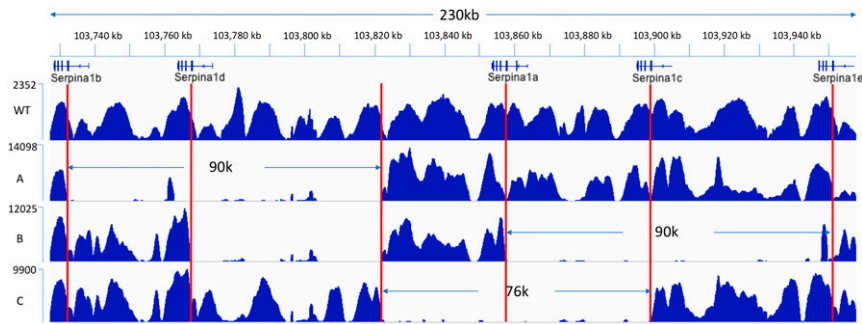


Fig. 2. Genomic characterization. A 230-kb genomic region was sequenced by PacBio-LITS for a wild-type C57BL/6J mouse and for a mouse from each founder line. Overview of the reads mapping to the 230-kb targeted region. (A) Line 7A; (B) line 24B; (C) line 31C. WT, wild type. Red lines indicate the location of the gRNAs.

F4 generation of line 24B, were compared with five C57BL/6 mice (all mice were 10-wk-old males) by RNA-Seq [single-end 50 base pair (bp) reads]. Founder #31 being female, the 31C line is at an earlier generation and was therefore not included in this analysis; however, preliminary data from a single animal do show a significant reduction of transcripts for all five paralogs (Dataset S1). RNA-Seq generated >50 million clean reads for each sample. A total of 231 genes were dysregulated in line 7A: 126 up-regulated genes and 105 down-regulated genes (Fig. 3A). A total of 299 genes were dysregulated in line 24B: 133 up-regulated genes and 166 down-regulated genes (Fig. 3B). The *Serpina* genes were the top down-regulated genes in both 7A and 24B lines (Datasets S2 and S3). The only exception to this was *Serpina1b* in line 24B. Reads mapping to *Serpina1b* were not decreased (fold change 0.9916; adjusted *P* value 0.9869). However, analysis of the reads revealed a deletion of two bases at position 314–315 of the mRNA, and additional deletions of four bases at position 347–350 and nine bases at position 351–359. These deletions overlap with the 73-bp genomic excision reported above, which corresponds to position 320–353 of the mRNA. Taken together with the genomic data, this suggests a combination of genomic excision events along gene disruptions resulting in unstable or truncated transcripts. In total, 106 genes were dysregulated in line 7A and line 24B jointly, 48 were up-regulated (marked in red, Fig. 3C) and 58 were down-regulated (marked in green, Fig. 3C and Dataset S4) including four of the *Serpina* genes (*a–c–d–e*). Pathway analysis revealed that the top dysregulated pathways were monocarboxylic acid and fatty acid metabolic processes, rhythmic process and circadian rhythm for line 7A (Fig. 3D), and rhythmic process and circadian rhythm for line 24B (Fig. 3E).

Off-targets for all four gRNAs were predicted using two independent algorithms: Cas OFFinder (16) and COSMID (17), and all results were compiled in Datasets S5 and S6. Results from these predictions were cross-referenced with the list of dysregulated genes (Datasets S2 and S3). No predicted off-target genes were identified as being dysregulated. Additionally, gene expression levels were manually checked for the predicted off-targets located within an ORF; no significant difference in gene expression was detected (Datasets S5 and S6). This supports the view that the described phenotype is not due to CRISPR off-targeting.

Induced Emphysema Following a Two-Hit, 2-Wk Lipopolysaccharide Challenge. Patients with AAT deficiency are subjected to lung damage when they are exposed to a bacterial infection. Bacterial infection leads to a neutrophilic response in the alveoli and the absence of AAT results in unopposed NE damage. The resulting emphysema leads to increased pulmonary compliance and decreased elastance. To determine whether the *Serpina1a–e* knockout mouse mimicked the clinical phenotype associated with this deficiency in humans, we exposed the airways of the mice to a mild lipopolysaccharide (LPS) challenge. The LPS challenge was

titrated to a dose that wild-type mice could easily overcome but could hypothetically lead to pathology in *Serpina1a–e* knockout mice due to the unabated activity of the secreted NE from the recruited neutrophils. For the LPS challenge, mice (*n* = 11–12 per group) received two doses of LPS on day 1 and day 12 before characterization of their pulmonary mechanics and lung morphometry at the end of the second week (Fig. 4A). At endpoint, pulmonary mechanics including maximal pressure–volume (PV) loops were assessed. The AAT knockout mice had evidence of

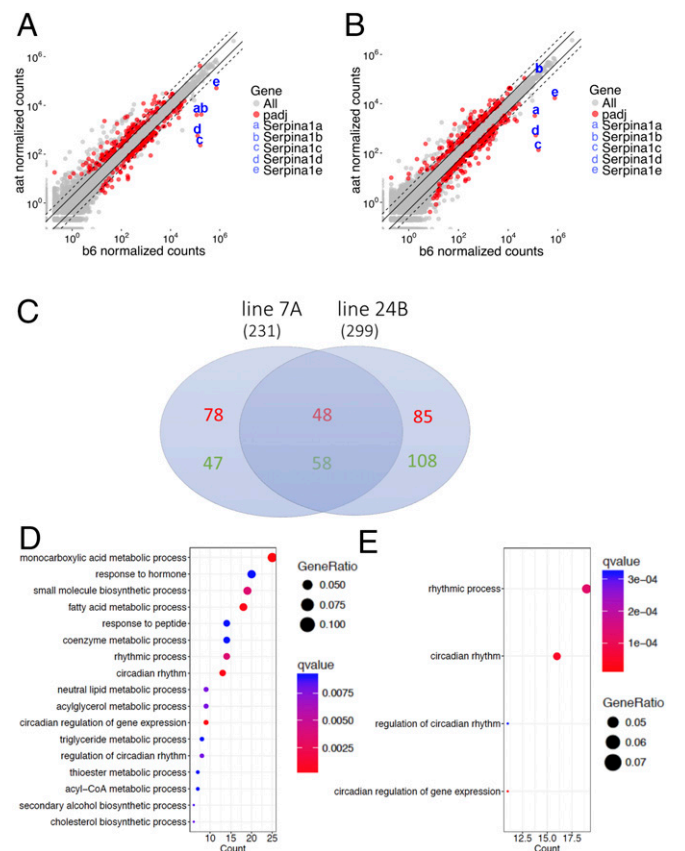


Fig. 3. Transcriptomic characterization. The transcriptome of five wild-type C57BL/6J mice, four mice from line 7A, and four mice from line 24B was characterized. The animals were aged and gender matched. Differential gene expression in knockout line vs. wild-type, for (A) line 7A, (B) line 24B. Solid line, fold change = 2; dashed line, fold change = 4. (C) Genes commonly dysregulated in lines 7A and 24B. Red, up-regulated genes; green, down-regulated genes. Pathway analysis for (D) line 7A and (E) line 24B.

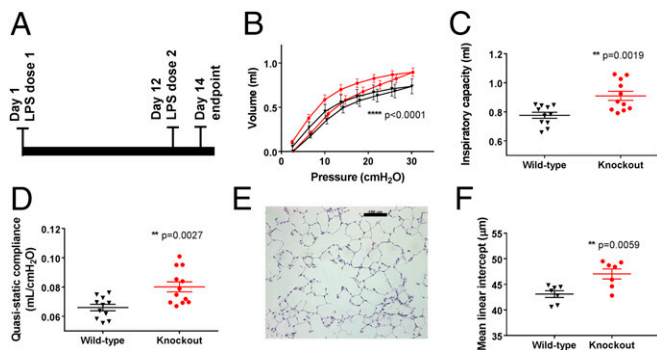


Fig. 4. Induced emphysema following a two-hit, 2-wk LPS challenge. (A) Wild-type mice (black triangles) and *Serpina1a-e* knockout mice (red circles) were treated twice with lipopolysaccharide (LPS) and presence of emphysema was assessed 2 wk after the first LPS dose. Pulmonary mechanics were determined using a flexiVent, parameters assessed include (B) pressure–volume (PV) loop, (C) inspiratory capacity, (D) quasistatic compliance (wild type $n = 11$, knockout $n = 12$). (E) Fixed lung tissue was stained by H&E, and a representative image from a wild-type mouse is shown here, which was then used for downstream image analysis. (Scale bar, 100 μm .) The morphometry was analyzed and the mean linear intercept quantified (F, $n = 7$ per group). The y axis is truncated in C, D, and F. Error bars represent the SEM. Statistical significance was determined by two-tailed unpaired t test, except for the PV loop (two-way ANOVA).

increased lung compliance as indicated by a PV loop that shifted up and to the left, as would be expected in an emphysematous lung (Fig. 4B). In addition, the inspiratory capacity of the knockouts was significantly increased from 0.78 mL to 0.91 mL (Fig. 4C). These changes were confirmed by measuring the quasistatic compliance which confirmed a significant increase from 0.066 mL/cmH₂O in the wild-type mice compared with 0.080 mL/cmH₂O in the AAT knockout mice (Fig. 4D). In summary, the pulmonary mechanics phenotype observed in the knockout mice is very comparable to that of patients with emphysema.

Next, alveolar morphometry, i.e., the quantitative assessment of alveolar architecture and structure, was determined. AAT-deficient patients present with panacinar emphysema, i.e., alveoli and alveolar ducts are uniformly enlarged (18). Therefore, a subset of the mice ($n = 7$ per group) was analyzed to assess their lung morphometry. Following pulmonary testing, the lungs were filled with a mixture of agarose and formalin to preserve their internal structure, lung volume was determined, and tissue sections were stained by H&E (example shown in Fig. 4E) before quantitative measurements described earlier (19). The total volume showed a trend toward the knockout group having on average a 22% higher volume (1.7 mL) than the wild-type group (1.4 mL, Fig. S4A) as well as a higher total surface area (knockout 149.2 μm^2 , wild type 130.8 μm^2 , Fig. S4B). The mean linear intercept, i.e., the mean free distance in the acinar air space complex, was significantly increased from 43.1 μm for the wild type to 47.05 μm for the knockouts (Fig. 4F). Taken altogether, these measurements indicate that in the *Serpina1a-e* knockouts, the alveolar ducts, and alveoli are enlarged, which recapitulates the clinical characteristics of emphysema.

Spontaneous Emphysema in Unchallenged, Aged *Serpina1a-e* Knockouts.

Given that a mild LPS challenge was able to recapitulate the hallmarks of emphysema, we sought to investigate whether emphysema would naturally develop in these *Serpina1a-e* knockout mice over time. Pulmonary mechanics were measured in age-matched, gender-matched wild-type and knockout mice at 35 (Fig. 5A and B) and 50 wk of age (Fig. 5C–F). At 35 wk of age, seven knockout and 10 wild-type animals were analyzed. The PV loop of the knockouts showed significantly increased compliance compared with their wild-type controls as evidenced by a shift upwards (Fig. 5A), which is characteristic of emphysema. Similarly, the quasistatic compliance is significantly increased in the knockouts compared with the wild types

(Fig. 5B). At 50 wk of age, nine knockout and 8 wild-type animals were analyzed. Again the PV loops are significantly different (Fig. 5C), and the difference between groups is larger at 50 wk than at 35 wk. The quasistatic compliance of the knockouts is also highly significantly increased from 0.0696 mL/cmH₂O in the wild types to 0.0989 mL/cmH₂O in the knockouts (Fig. 5D, $P \leq 0.0001$), and the elastance is significantly decreased (Fig. 5E, $P \leq 0.0002$). The resistance is similar in the wild-type controls and the knockouts (Fig. 5F). Taken altogether, these results indicate that the *Serpina1a-e* knockout mice spontaneously develop emphysema, with early signs detectable at 35 wk of age, and at 50 wk of age they present with a clear emphysematous phenotype.

Cellular and Histological Changes in the Lungs of *Serpina1a-e* Knockouts.

To determine the mechanism driving the spontaneous emphysema, we analyzed the lung bronchoalveolar lavage fluid (BALF) of aging mice to see whether there was an active mechanism or process that could explain the progression of lung disease. We hypothesized that in the absence of an exogenous insult such as the LPS or a cigarette-smoke challenge the *Serpina1a-e* knockout mice must have an autonomous increase in the protease/antiprotease imbalance in the lung despite being housed in a specific pathogen-free (SPF) facility.

Cell counts were determined in the bronchoalveolar lavage (BAL) of wild-type and *Serpina1a-e* knockout mice at 8 and 42 wk of age. Consistent with our hypothesis, neutrophil, monocyte, lymphocyte, and macrophage counts were all increased in the 42-wk-old knockout mice (Fig. 6A), while no difference was detectable in the 8-wk-old mice. We then measured elastase activity in the BALF, to determine whether the increase in the cell infiltration was sufficient to increase the protease/antiprotease imbalance (Fig. 6B). These

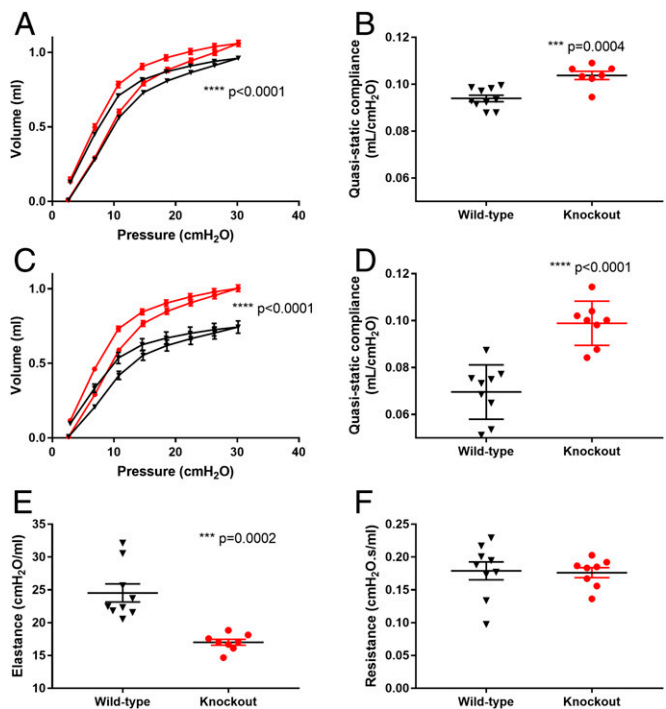


Fig. 5. Spontaneous emphysema in unchallenged, aged *Serpina1a-e* knockouts. Pulmonary mechanics were assessed using a flexiVent in 35-wk-old (A and B, wild type $n = 10$, knockout $n = 7$) and 50-wk-old (C–F, wild type $n = 9$, knockout $n = 8$) wild-type mice (black triangles) and *Serpina1* knockout mice (red circles). Key parameters assessed include pressure–volume (PV) loop (A and C) and quasistatic compliance (B and D). Additional parameters include elastance (E) and resistance (F). The y axis is truncated in B, D, and E. Error bars represent the SEM. Statistical significance was determined by two-tailed unpaired t test, except for the PV loops (two-way ANOVA).

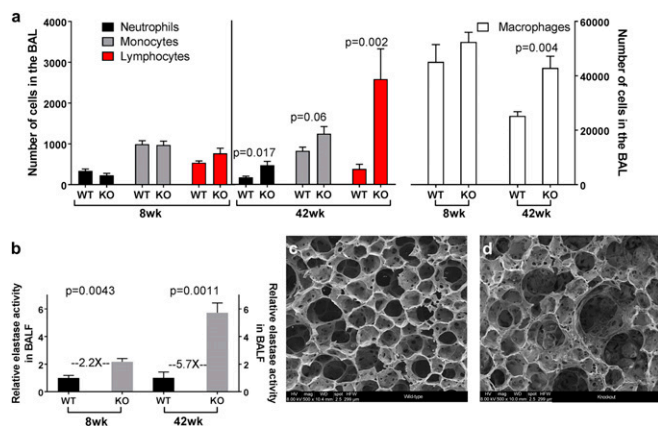


Fig. 6. Cellular and histological changes in the lungs of *Serpina1a-e* knockouts. (A) Cell counts in the BAL (bronchoalveolar lavage) including neutrophils (black), monocytes (gray), lymphocytes (red), and macrophages (white) were determined at 8 wk and 42 wk of age ($n = 5$ per group). (B) Relative elastase activity in bronchoalveolar lavage fluid (BALF) was quantified at 8 wk and 42 wk of age ($n = 5$ per group). Alveolar structure was determined at 70 wk of age by scanning electron microscopy (C and D).

results also indicate an increase in elastase activity with age, where in the 8-wk cohort, there was a 2.2-fold difference between wild-type and knockout groups ($P < 0.0043$), and this difference was increased to 5.7-fold in the 42-wk cohort ($P < 0.0011$). Finally, scanning electron microscopy was performed on the fixed lungs of 70-wk-old mice to confirm the structural changes of the alveoli as a result of the spontaneous age-dependent emphysema. The images obtained demonstrate enlargement of the alveolar ducts and alveoli in *Serpina1a-e* knockouts (Fig. 6 C and D).

Discussion

The *Serpina1a-e* null mouse model described in this report was generated by disrupting five highly homologous gene paralogs through the targeting of exon 2 using a combination of four gRNAs. This was achieved three times independently to give rise to three mouse lines. The genes were disrupted in some cases by creating large genomic excisions of up to 90,000 bp, but in some other cases the genomic alterations were more difficult to pinpoint, due to the very high degree of homology of the paralogs. The null status of the mouse lines was therefore confirmed by transcriptome analysis, as NMD would occur when a frameshift-causing alteration of exon 2 would be created. Indeed in lines A and C, transcripts were highly significantly reduced; however, in line B transcripts were detected at unchanged levels for *Serpina1b*. It is unclear why NMD does not occur in this instance. The genomic sequence of *Serpina1b* in line B was precisely characterized, it is a 73-bp excision linked to gRNA#2, which indeed induces a frameshift. Translational predictions indicate the potential generation of a truncated, 105-residue long protein.

The *Serpina1a-e* null mouse model presented here was generated by zygote injections using CRISPR/Cas9 targeting, which resulted in three founders with a biallelic quintuple gene knockout, highlighting the efficiency and robustness of this technology. This characterization is supported by genomic, transcriptomic, protein, and physiological data. The animals spontaneously develop emphysema with age, as shown by pulmonary mechanics, a phenotype consistently observed in all generations of three independent lines. Here we report that the mechanism underlying this phenotype is due to an increase in the cellular infiltrates of these mice as they age, which leads to an increase in the elastase activity in the BAL. This in turn results in increased alveolar wall and elastin fiber degradation in the face of a lower antielastase capacity in the *Serpina1a-e* knockout mice. It is possible that in the absence of AAT, the elastin degradation products lead to further recruitment of cells into the lungs and amplifies this

degenerative process in a feedforward loop that leads to irreversible lung damage over time. This increased recruitment of cellular infiltrates in AAT-deficient lungs is supported by evidence documenting the consequence of increased pulmonary neutrophil elastase activity (20, 21) and of elastin breakdown products in AAT deficiency (22–24). This phenotype models α -1 antitrypsin deficiency lung disease, and should prove extremely valuable to the preclinical development of novel therapeutics for this disease that affects about 100,000 individuals in the United States and millions worldwide. This unique model may allow a precise dissection of the disease mechanism and bring further insight into the genetic pathways linked to the pathology. Additionally, the current clinical endpoint for therapeutic candidates is serum protein levels of 11 μ M, a threshold defined through genotype/phenotype studies. However, studies using this animal model may allow further refinement of this threshold, and the definition of alternate endpoints. Finally, this murine strain constitutes the ideal genetic background in which to introduce the most common human mutation associated with α -1 antitrypsin deficiency, the Z form of the protein (Z-AAT). Currently ongoing, this work will result in a mouse model that will genetically and phenotypically recapitulate two components of the human disease: a loss of function in the lungs and a gain of function in the liver. The generation of this model will support the preclinical development of genome editing therapeutics for α -1 antitrypsin deficiency.

It is well established that patients with α -1 antitrypsin deficiency are more likely to develop emphysema and/or develop it earlier if they are exposed to environmental inhalants, which include smoking (active or passive) and occupational exposure (25–27). This first animal model of α -1 antitrypsin deficiency will therefore be highly relevant to further investigate those findings. Beyond the scope of α -1 antitrypsin, this mouse model more broadly constitutes a unique model in which to study emphysema. It is a genetic model of emphysema, not one based on developmental impairment of alveolarization, or reliant upon administration of elastase or other proteinases. It may therefore be used to test factors that influence development of postnatal or age-related emphysema. For example, the model might enable the investigation of occupational exposure to untested air particulates, smoking, and smoke exposure, as well as new behaviors such as vaping. In conclusion, this mouse model may have a significantly impact on preventive health for the α -1 antitrypsin patient community, but also for the COPD population at large, and constitutes a significant step forward for COPD research.

Materials and Methods

Animal Experiments. The institutional animal care and use committee at the University of Massachusetts Medical School approved all animal experiments. Further information is provided in *SI Materials and Methods*.

Design and Validation of Guide RNAs. National Center for Biotechnology Information (NCBI) blast search was used to identify conserved regions within the mouse *Serpina1* paralogs. CRISPR/Cas9 target sites present in all isoforms and with low predicted off-target activity were identified using the Bio-conductor software package CRISPRseek (28). Four CRISPR/Cas9 target sequences were selected that are conserved in each of the mouse *Serpina* paralogs.

ELISA. AAT levels in serum samples were quantified by direct ELISA as described (29). Briefly, serum proteins were immobilized in each well, and AAT was detected by incubating with a polyclonal goat anti-mouse AAT antibody (GA1T-90A-Z, ICL), followed by incubation with an HRP-conjugated rabbit anti-goat antibody (ICL) and a peroxidase substrate (KPL Scientific). OD_{450nm} was read and a four-parameter logistic model was fit to the standard curve (RS-90A1T, ICL); subsequently, AAT concentration of the unknown samples was derived. Samples were run in technical triplicates.

Targeted Genome Sequencing. A 230-kb region of chromosome 12 (chr12:103,726,970–103,959,013) was sequenced by PacBio-large insert targeted sequencing (PacBio-LITS), following a protocol adapted from ref. 30. Briefly, high molecular weight gDNA was isolated from fresh liver tissue of adult mice (Genomic-tip 500, Qiagen), sheared to ~ 6 kb (g-TUBE, Covaris), preamplified,

hybridized with custom probes (Roche Nimblegen) designed based on mm10/GRCh38 for target enrichment by bead capture, and the captured libraries were amplified. Finally, SMRT libraries were constructed and sequenced on a PacBio instrument at the Yale Center for Genome Analysis. Raw reads were broken into separate subreads after removing adapters and trimming away low-quality regions. PacBio sequencing generated >640,000 subreads for each sample. Subreads were aligned to the 230-kb target region by Burrows–Wheeler aligner (BWA)-maximal exact matches. Using Canu (31), all the reads with high coverage (>30) are assembled locally into contigs.

Additionally, TLA was performed on line B as described earlier (32). Splenocytes were isolated from a wild-type mouse and a knockout mouse from line B and sent to Cergentis B.V. for TLA analysis. Briefly, DNA was cross-linked, fragmented, and religated. Primer pairs complementary to sequences of each *Serpina1* paralog were then used to amplify the surrounding loci. PCR products were purified and library was prepped using the Illumina NexteraXT protocol and sequenced on an Illumina sequencer. Reads were mapped using BWA's Smith–Waterman (SW) alignment tool and mouse mm10 genome version.

RNA-Seq. Total RNA was isolated from snap-frozen liver tissue of 10-wk-old male mice using TRIzol (Life Technologies) following the manufacturer's instructions. Barcoded libraries were sequenced as 50-nt single end reads on an Illumina HiSeq 4000 instrument at the Beijing Genomics Institute. After

sequencing, the raw reads were filtered, which includes removing adapter sequences, contamination, and low-quality reads. RNA-Seq generated >50 million clean reads for each sample. Using TopHat2 (33), at least 70% of the clean reads mapped at a single location of the mouse reference genome (mm10) for each replicate sample. Normalization and differential expression gene analysis were performed with HTSeq-count (34) and DESeq2 (35) software packages. Pathway analysis was performed as described in ref. 36.

FlexiVent. Measurements of pulmonary mechanics were performed using forced oscillometry (flexiVent system; SCIREQ). Respiratory mechanics were obtained and calculated using flexiWare software (SCIREQ) as previously described (37).

Morphometry. Immediately following measurements of lung mechanics (flexiVent), lungs were harvested for morphometric analysis as previously described (19, 38).

ACKNOWLEDGMENTS. We acknowledge the support of the University of Massachusetts Transgenic Animal Modeling Core and Mutagenesis Core. This work was funded by a University of Massachusetts Medical School startup grant (to C.M.); NIH Grants R01-NS088689 (to C.M.), R01-DK098252 (to C.M. and T.R.F.), and R24-OD018259 (to C.M.); and the Alpha-1 Foundation (C.M.). F.B. is supported by an Alpha-1 Foundation Postdoctoral Fellowship.

- Lozano R, et al. (2012) Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the global burden of disease study 2010. *Lancet* 380:2095–2128.
- Buist AS, et al.; BOLD Collaborative Research Group (2007) International variation in the prevalence of COPD (the BOLD study): A population-based prevalence study. *Lancet* 370:741–750.
- Lomas DA, Silverman EK (2001) The genetics of chronic obstructive pulmonary disease. *Respir Res* 2:20–26.
- Laurell CB, Eriksson S (2013) The electrophoretic α 1-globulin pattern of serum in α 1-antitrypsin deficiency. 1963. *COPD* 10:3–8.
- Flotte TR, et al. (2011) Phase 2 clinical trial of a recombinant adeno-associated viral vector expressing alpha(1)-antitrypsin: Interim results. *Hum Gene Ther* 22:1239–1247.
- Mueller C, et al. (2013) Human Treg responses allow sustained recombinant adeno-associated virus-mediated transgene expression. *J Clin Invest* 123:5310–5318.
- Mueller C, et al. (2017) 5 year expression and neutrophil defect repair after gene therapy in alpha-1 antitrypsin deficiency. *Mol Ther* 25:1387–1394.
- Mueller C, et al. (2012) Sustained miRNA-mediated knockdown of mutant AAT with simultaneous augmentation of wild-type AAT has minimal effect on global liver miRNA profiles. *Mol Ther* 20:590–600.
- Borel F, et al. (2017) Survival advantage of both human hepatocyte xenografts and genome-edited hepatocytes for treatment of alpha-1 antitrypsin deficiency. *Mol Ther* 25:2477–2489.
- Takubo Y, et al. (2002) Alpha1-antitrypsin determines the pattern of emphysema and function in tobacco smoke-exposed mice: Parallels with human disease. *Am J Respir Crit Care Med* 166:1596–1603.
- Barbour KW, et al. (2002) The murine alpha(1)-proteinase inhibitor gene family: Polymorphism, chromosomal location, and structure. *Genomics* 80:515–522.
- Wang D, et al. (2011) Deletion of *Serpina1a*, a murine α 1-antitrypsin ortholog, results in embryonic lethality. *Exp Lung Res* 37:291–300.
- Kushi A, et al. (2004) Disruption of the murine alpha1-antitrypsin/PiZ gene. *Exp Anim* 53:437–443.
- Paterson T, Moore S (1996) The expression and characterization of five recombinant murine alpha 1-protase inhibitor proteins. *Biochem Biophys Res Commun* 219:64–69.
- Yang H, Wang H, Jaenisch R (2014) Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat Protoc* 9:1956–1968.
- Bae S, Park J, Kim JS (2014) Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30:1473–1475.
- Cradick TJ, Qiu P, Lee CM, Fine EJ, Bao G (2014) COSMID: A web-based tool for identifying and validating CRISPR/Cas off-target sites. *Mol Ther Nucleic Acids* 3:e214.
- Takahashi M, et al. (2008) Imaging of pulmonary emphysema: A pictorial review. *Int J Chron Obstruct Pulmon Dis* 3:193–204.
- Paxson JA, et al. (2011) Age-dependent decline in mouse lung regeneration with loss of lung fibroblast clonogenicity and increased myofibroblastic differentiation. *PLoS One* 6:e23232.
- Hubbard RC, et al. (1991) Neutrophil accumulation in the lung in alpha 1-antitrypsin deficiency. Spontaneous release of leukotriene B4 by alveolar macrophages. *J Clin Invest* 88:891–897.
- Nakamura H, Yoshimura K, McElvaney NG, Crystal RG (1992) Neutrophil elastase in respiratory epithelial lining fluid of individuals with cystic fibrosis induces interleukin-8 gene expression in a human bronchial epithelial cell line. *J Clin Invest* 89:1478–1484.
- Houghton AM, et al. (2006) Elastin fragments drive disease progression in a murine model of emphysema. *J Clin Invest* 116:753–759.
- Hunninghake GW, et al. (1981) Elastin fragments attract macrophage precursors to diseased sites in pulmonary emphysema. *Science* 212:925–927.
- Senior RM, Griffin GL, Mecham RP (1980) Chemotactic activity of elastin-derived peptides. *J Clin Invest* 66:859–862.
- Mayer AS, et al. (2000) Occupational exposure risks in individuals with Pi*Z alpha(1)-antitrypsin deficiency. *Am J Respir Crit Care Med* 162:553–558.
- Piittala E, Tornling G, Eriksson S (1997) Effect of age and occupational exposure to airway irritants on lung function in non-smoking individuals with alpha 1-antitrypsin deficiency (PiZZ). *Thorax* 52:244–248.
- Senn O, Russi EW, Imboden M, Probst-Hensch NM (2005) alpha1-Antitrypsin deficiency and lung disease: Risk modification by occupational and environmental inhalants. *Eur Respir J* 26:909–917.
- Zhu LJ, Holmes BR, Aronin N, Brodsky MH (2014) CRISPRseek: A bioconductor package to identify target-specific guide RNAs for CRISPR-Cas9 genome-editing systems. *PLoS One* 9:e108424.
- Cox A, Mueller C (2017) Quantification of murine AAT by direct ELISA. *Methods Mol Biol* 1639:217–222.
- Wang M, et al. (2015) PacBio-LITS: A large-insert targeted sequencing method for characterization of human disease-associated chromosomal structural variations. *BMC Genomics* 16:214.
- Koren S, et al. (2017) Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27:722–736.
- Hottentot QP, van Min M, Splinter E, White SJ (2017) Targeted locus amplification and Next-generation sequencing. *Methods Mol Biol* 1492:185–196.
- Kim D, et al. (2013) TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36.
- Anders S, Pyl PT, Huber W (2015) HTSeq—A python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166–169.
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550.
- Yu G, Wang LG, Han Y, He QY (2012) clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* 16:284–287.
- McGovern TK, Robichaud A, Fereydoonzad L, Schuessler TF, Martin JG (2013) Evaluation of respiratory system mechanics in mice using the forced oscillation technique. *J Vis Exp* e50172.
- Davis AM, Thane KE, Hoffman AM (2017) Practical methods for assessing emphysema severity based on estimation of linear mean intercept (Lm) in the context of animal models of alpha-1 antitrypsin deficiency. *Methods Mol Biol* 1639:93–106.