# The Ser82 RAGE Variant Affects Lung Function and Serum RAGE in Smokers and sRAGE Production *In Vitro*

Suzanne Miller[1]\*, Amanda P. Henry[1], Emily Hodge[1], Alexander K. Kheirallah[1], Charlotte K. Billington[1], Tracy L. Rimington[1], Sangita K. Bhaker[1], Ma'en Obeidat[1¤a], Erik Melén[2¤b], Simon K. Merid[2], Caroline Swan[3], Catherine Gowland[1], Carl P. Nelson[1], Ceri E. Stewart[1], Charlotte E. Bolton[1], Iain Kilty[4], Anders Malarstig[5], Stuart G. Parker[6], Miriam F. Moffatt[7], Andrew J. Wardlaw[8], Ian P. Hall[1], Ian Sayers[1]

**1** Division of Respiratory Medicine, University of Nottingham, Nottingham, United Kingdom, **2** Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden, **3** Department of Biology, University of York, York, United Kingdom, **4** Pfizer Worldwide Research & Development, Cambridge, Massachusetts, United States of America, **5** Pfizer Worldwide Research & Development, Cambridge, United Kingdom, **6** Institute for Ageing and Health, Newcastle University, Campus for Ageing and Vitality, Newcastle upon Tyne, United Kingdom, **7** National Heart and Lung Institute, Imperial College London, London, United Kingdom, **8** Institute for Lung Health, Immunity and Inflammation, University of Leicester, Leicester, United Kingdom

¤a Current address: James Hogg Research Centre, Institute for Heart and Lung Health, University of British Columbia, Vancouver, British Columbia, Canada
¤b Current address: Sachs' Children's Hospital, Stockholm, Sweden
\* Suzanne.Miller@nottingham.ac.uk

## Abstract

### Introduction

Genome-Wide Association Studies have identified associations between lung function measures and Chronic Obstructive Pulmonary Disease (COPD) and chromosome region 6p21 containing the gene for the Advanced Glycation End Product Receptor (*AGER*, encoding RAGE). We aimed to (i) characterise RAGE expression in the lung, (ii) identify *AGER* transcripts, (iii) ascertain if SNP rs2070600 (Gly82Ser C/T) is associated with lung function and serum sRAGE levels and (iv) identify whether the Gly82Ser variant is functionally important in altering sRAGE levels in an airway epithelial cell model.

### Methods

Immunohistochemistry was used to identify RAGE protein expression in 26 human tissues and qPCR was used to quantify *AGER* mRNA in lung cells. Gene expression array data was used to identify *AGER* expression during lung development in 38 fetal lung samples. RNA-Seq was used to identify *AGER* transcripts in lung cells. sRAGE levels were assessed in cells and patient serum by ELISA. BEAS2B-R1 cells were transfected to overexpress RAGE protein with either the Gly82 or Ser82 variant and sRAGE levels identified.

## Results

Immunohistochemical assessment of 6 adult lung samples identified high RAGE expression in the alveoli of healthy adults and individuals with COPD. *AGER*/RAGE expression increased across developmental stages in human fetal lung at both the mRNA (38 samples) and protein levels (20 samples). Extensive *AGER* splicing was identified. The rs2070600T (Ser82) allele is associated with higher $FEV_1$, $FEV_1$/FVC and lower serum sRAGE levels in UK smokers. Using an airway epithelium model overexpressing the Gly82 or Ser82 variants we found that HMGB1 activation of the RAGE-Ser82 receptor results in lower sRAGE production.

## Conclusions

This study provides new information regarding the expression profile and potential role of RAGE in the human lung and shows a functional role of the Gly82Ser variant. These findings advance our understanding of the potential mechanisms underlying COPD particularly for carriers of this *AGER* polymorphism.

## Introduction

Lung diseases including chronic obstructive pulmonary disease (COPD) remain a major burden to society and the economy. In recent years there has been substantial progress in our understanding of genetic factors associated with lung function parameters through genome wide association approaches [1–4]. There is a key need however, to define the mechanisms which underlie the identified associations. Meta-analyses of genome-wide association studies (GWAS) have identified polymorphisms in the Advanced Glycation End Product-Specific gene (*AGER*, pivotal SNP in exon 3 rs2070600 (Gly82Ser, C/T)) as being associated with Forced Expiratory Volume in 1 second ($FEV_1$)/Forced Vital capacity (FVC) [1, 2]. The role of *AGER* now requires further elucidation in the lung which may provide new insight into the mechanisms underlying several lung diseases where *AGER* genetic variants have been implicated, including e.g. COPD and lung cancer [5, 6]. *AGER* encodes RAGE; a member of the immunoglobulin superfamily which interacts with a diverse range of ligands, including damage associated molecular pattern proteins such as High mobility group protein B1 (HMGB1). Activation of RAGE has many downstream effects including the regulation of inflammation, cell migration, proliferation and adhesion [7, 8]. Importantly, accumulating data suggests RAGE may play a role in a range of diseases including COPD [5, 9, 10], Alzheimer's disease, arthritis and diabetes [11–13] and so a greater understanding of *AGER* regulation by genetic mechanisms has broad implications. RAGE is expressed abundantly in the membrane and cytoplasm of both pneumocytes and macrophages in lung [14], suggesting an important role in lung homeostasis. Recent studies in transgenic mouse models have shown a role for RAGE in alveolar morphogenesis during lung development [15] with overexpression of RAGE also being shown to result in the development of an emphysema-like phenotype in adult mice [16]. Increased expression of RAGE, or the soluble forms of the receptor generated by cleavage (sRAGE) or mRNA splicing (esRAGE), has been found in the airways and serum of COPD patients [5, 9, 10]. However, the mechanism underlying the genetic association with lung function seen at this locus has as yet not been defined. In the current study we set out specifically to further define the potential role of *AGER* in the context of respiratory disease. To do this we

have determined the expression profile of RAGE in human lung and during lung development. Additionally, we have investigated the specific functional effects of the key Gly82Ser variant that has been associated with numerous lung phenotypes at both the population and cellular levels providing an insight into the function of this variant in the lung.

## Methods

### Immunohistochemistry

Three control (no history of obstructive lung disease) adult lung samples and three samples from individuals with clinically diagnosed COPD were provided by the Nottingham Health Science Biobank following isolation by lung resection (Nottingham, UK) under ethical approval (08/H0407/1). Twenty fetal lung tissue samples were provided by The Human Developmental Biology Resource (www.hdbr.org) at different stages, specifically 19, 21 and 23 days and 9, 10, 11, 12, 13, 14, 15, 16, 17 and 19 weeks post-conception. Immunohistochemistry was performed as previously described [17]. Slides were incubated with either a rabbit polyclonal anti-RAGE antibody (1:500, ab30381, Abcam, Cambridge, UK) or normal rabbit IgG isotype control (Invitrogen/Life Technologies Ltd.) for 1 hour at room temperature.

### Affymetrix U133 Plus 2 array data for human fetal lung

Prior studies have obtained human fetal lung tissues from the National Institute of Child Health and Human Development tissue databases and the extracted RNA profiled using microarrays to investigate gene expression spanning different gestational ages [18, 19]. RNA samples from 38 subjects (estimated gestational age 7–22 weeks or 53–154 days post conception) *i.e.* Pseudoglandular (gestational age, 7–16 weeks) and into the Canalicular (17–26 weeks) stages of development were used in the current *AGER* analyses. These data are available at NCBI Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo), GSE14334. Affymetrix probes; 217046_s_at and 210081_at were used for *AGER* expression analyses in relation to gestational age (linear regression model (lmFit) as implemented in the Limma package in R/Bioconductor).

### Quantitative PCR

Quantification of total *AGER* mRNA was carried out using total RNA from cultured cells or commercially available total lung tissue (Ambion/Life Technologies) using Taqman™ (Applied Biosystems/Life Technologies) as previously described [20]. In each sample, quantitative PCR (qPCR) was carried out for *AGER* (Assay VY4331182 designed to detect part of the 5'region and thus detects the majority of transcripts, Invitrogen) and Human 18S ribosomal RNA 1 endogenous control (4310893E, Applied Biosystems). qPCR was performed using a Stratagene Mx3005P machine (Agilent Technologies, UK). Data were analysed using the $2^{-\Delta\Delta Ct}$ method.

### RNA Sequencing

Total RNA was extracted from normal human bronchial epithelial cells (NHBECs) from three donors (Lonza, UK) at either passage 3 or 4 using the GenElute Mammalian Total RNA Miniprep Kit (Sigma-Aldrich). Donor 1 was a Caucasian, 56 years old, Male, Smoker; Donor 2 was a 19 year old, Male, Smoker and Donor 3 was a Caucasian, 50 year old, Male, Smoker. RNA quality was assessed using the Agilent 2100 Bioanalyzer with all samples having a RNA integrity number $\geq$ 8. A sequencing library was prepared using TruSeq RNA Sample Prep Kit v2 (Illumina). mRNA was poly-A selected using oligo-dT coated magnetic beads and fragmented. cDNA was synthesised using random primers. Finally, a cDNA library was prepared by end-

repair, phosphorylation, A-tailing, adapter ligation and PCR amplification. Paired-end sequencing was performed on either the Illumina HiSeq 2000 or the NextSeq 500 with approximately 40 million reads per sample generated. RNA-seq data was analysed as described [21]. The raw read FastQ files (100 base pair reads) were quality evaluated using FastQC. Median quality scores were above 28 for all samples. Reads were used for subsequent analysis using the Ubuntu 12.04 LTS operating system. Reads were aligned to the human genome (Build GRCh37) for each sample using Bowtie2 tool as part of TopHat (v2.0.12). Reads aligning to more than 20 positions were discarded. The Cufflinks v2.2.1 programme was used to assemble transcriptomes for each individual sample. Transcriptomes from all the samples were merged by reference annotation based transcript assembly (RABTA) using Cuffmerge v1.0.0 to provide the maximum reads to identify low-expression transcripts. The Cuffmerge generated novel gene transfer format (GTF) annotation file was compared to the Ensembl GTF annotation of the GRCh37 genome build using Cuffcompare v2.2.1. Basal NHBEC sample data (n = 3) were used to identify transcript abundance. Mean fragments per kilobase of transcript per million fragments mapped (FPKM) expression of isoforms was calculated and percent of total transcripts determined. Splicing graphs were generated using SpliceGrapher v0.2.4.

## Patient recruitment, genotyping and ELISA

Subjects were recruited from five UK centres for smoking history and/or COPD diagnosis, this cohort is the same as that previously described [22, 23]. This generated a cohort for lung function association analyses composed of 1042 smokers with DNA samples available. rs2070600 was genotyped using the LGC Genomics KBioscience service (Herts, UK). We generated a subgroup based on rs2070600 genotype composed of 102 subjects for serum analyses (51 C:T and 51 C:C age, sex and smoking pack year matched). The level of sRAGE in the serum was determined using the Quantikine ELISA kit for Human RAGE as directed by the manufacturer (R & D systems, UK). Ethical approval was obtained from local ethics committees (Nottingham, Sheffield, Manchester, Leicester and Oxford).

## Site-directed mutagenesis and transfection of BEAS2B-R1 cells

A plasmid containing the cDNA for the full length, membrane form of RAGE (pcDNA3.1-fl-RAGE) was a kind gift of Dr. B. Hudson (University of Miami). Site directed mutagenesis of the flRAGE plasmid was carried out using QuikChange® Multi Site-directed Mutagenesis kit (Agilent, UK). A single mutagenic primer (CGTGTCCTTCCCAAC**(A)**GCTCCCTCTTCCT) was used to generate the alternative allele of SNP rs2070600. All constructs were sequence verified (data not shown). BEAS2B-R1 epithelial cells (provided by Dr. Ray Penn, University of Maryland, Baltimore, USA) [24] were transiently transfected with plasmids using Fugene 6 (Promega, UK) at a 3:2 Fugene:DNA ratio, grown for 24 hours and then serum starved for 16 hours. 250ng/ml of HMGB1 ligand (R & D systems, UK, 1690-HMB-025) was then added to the cells for 48 hours. Supernatants were collected for analysis and sRAGE levels measured as described for serum analyses. Cells were lysed and RNA extracted using the GenElute Mammalian Total RNA miniprep kit (Sigma). cDNA was synthesised using the Superscript First-Strand Synthesis System for RT-PCR (Invitrogen/Life Technologies). qPCR for *AGER* was performed as described. Data from four independent cell experiments were combined for analyses.

## Statistical analysis

Linear regression was used to assess association between lung function measures ($FEV_1$, FVC and $FEV_1$/FVC) and rs2070600 genotype or log serum sRAGE levels as previously described [2]. Mann-Whitney tests were used to compare groups (GraphPad Prism 6).

## Results

### RAGE expression and localisation in healthy and COPD lungs

The expression of RAGE (encoded by *AGER*) was determined in the alveoli and bronchi of control and COPD lungs using immunohistochemistry. Specific staining was observed in both the cytoplasm and membrane (Fig 1). Pneumocytes within the alveolar regions of all three lung samples from control individuals exhibited strong staining for RAGE (Fig 1a) and in these samples the bronchial epithelium showed a variable, weaker staining pattern for RAGE (Fig 1e). In the lung tissue of all three individuals with COPD, strong immuno-positivity for
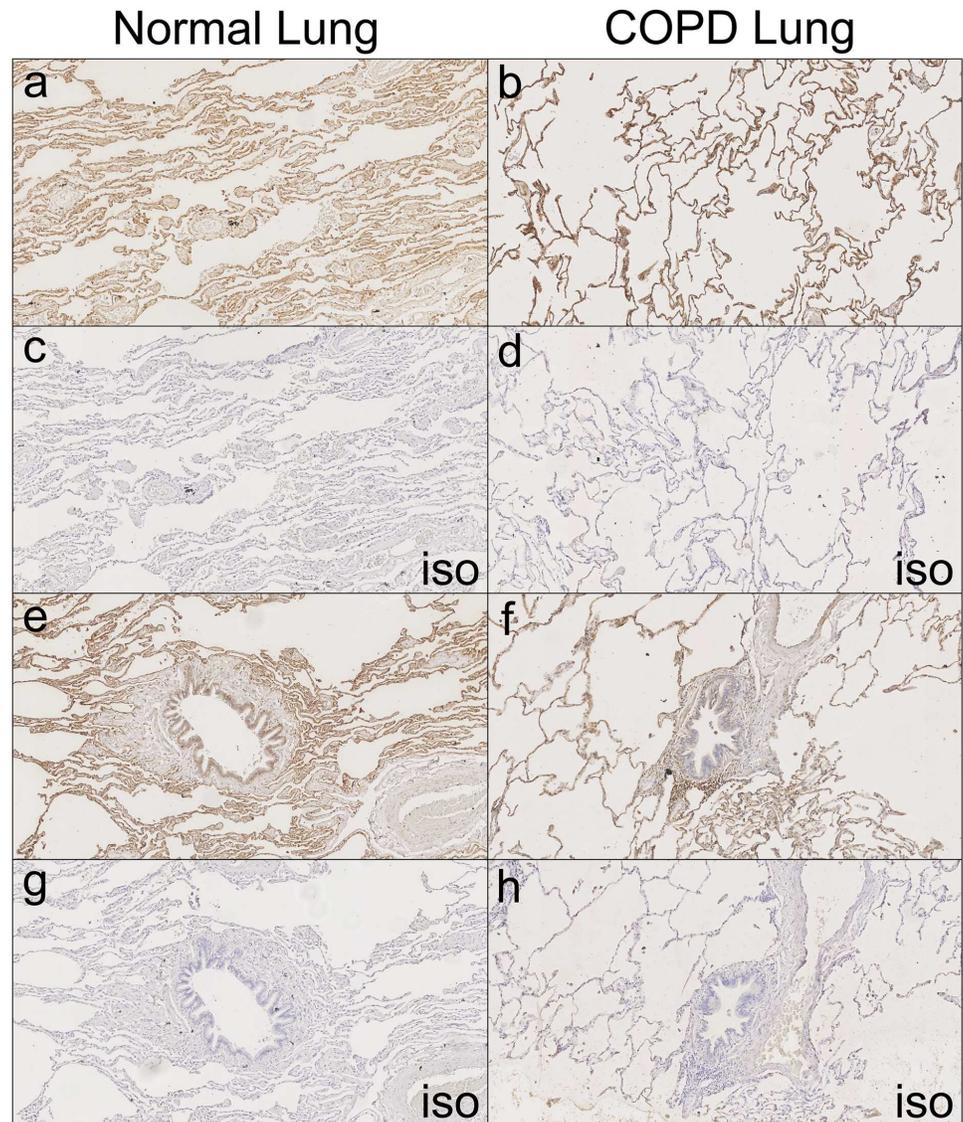


**Fig 1. Immunohistochemical analysis of RAGE expression in healthy and COPD lung.** In healthy lung tissue, RAGE was found to be localised to the cytoplasm and membrane. RAGE expression was high in the pneumocytes of alveolar regions (a). The bronchial epithelium showed variable weak to moderate staining (e). In lung tissue of individuals with COPD, RAGE was very strongly immunopositive in the membrane and cytoplasm of the pneumocytes in the alveolar regions (b). The bronchial epithelium from individuals with COPD was weak or immunonegative for the RAGE protein (f). All isotype controls were negative (c, d, g and h). Representative images of one healthy and one COPD lung shown. x10 magnification.

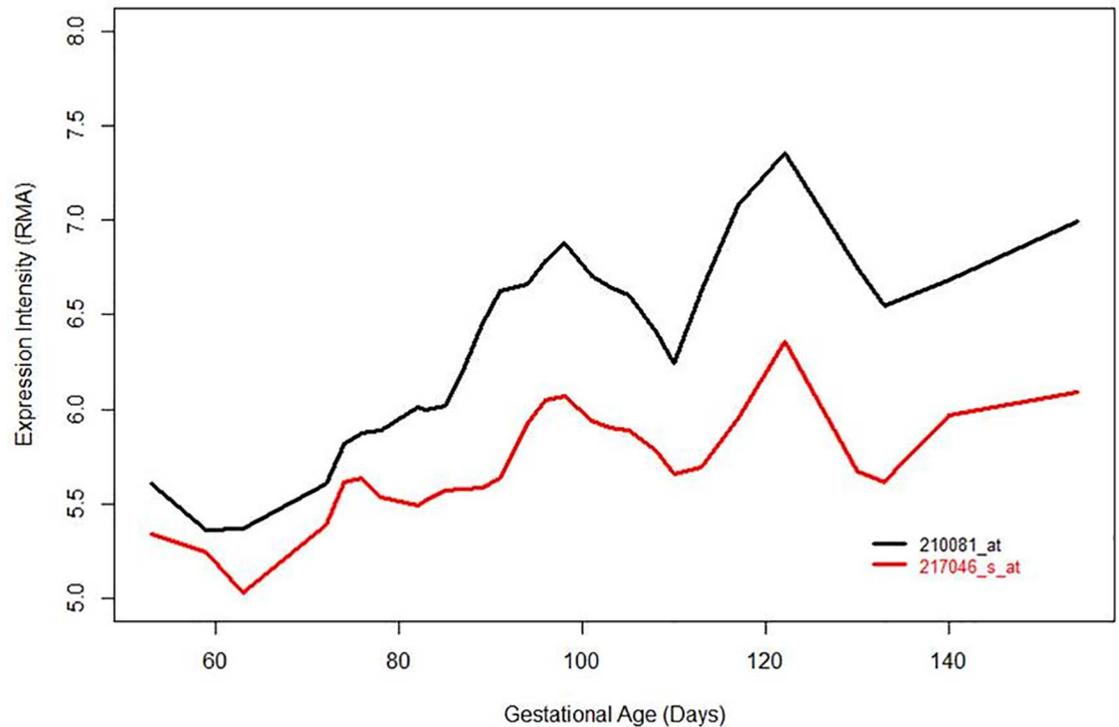doi:10.1371/journal.pone.0164041.g001

**Fig 2. Gene expression array data of 38 fetal lungs for the *AGER* probes.** Expression intensities for the *AGER* probes 210081_at and 217046_s_at were plotted against the gestational age of each fetal lung sample and showed an increase in expression with fetal lung age. RMA; Robust Multi-array Average. Affymetrix U133 Plus 2 expression array probe 210081_at probes for exon 8 of AGER mRNA whilst 217046_s_at probes AGER exons 7–11'.

doi:10.1371/journal.pone.0164041.g002

RAGE was found both in the membrane and cytoplasm of pneumoctyes in the alveolar regions (Fig 1b), whilst the bronchial epithelium showed limited RAGE expression (Fig 1f).

## AGER mRNA and RAGE protein expression increase during lung development

The abundance of *AGER* mRNA in thirty eight human fetal lung samples was assessed using publically available expression microarray data (see methods). Levels of *AGER* mRNA were found to increase with gestational age across the Pseudoglandular and into the Canalicular stages of human lung development (gestational ages, 7–16 weeks and 17–26 weeks, respectively), (Fig 2 and Table 1).

  38 human lung tissue samples from gestational ages 7–22 weeks (or 53–154 days post conception) were assessed. AveExpr = average expression between all samples, t = t-statistic describing differential expression, *P* value = unadjusted *P* value, Adj.P.value = adjusted *P* value controlling for false discovery rate, Beta-coefficient = log (odds ratio) corresponding to the

**Table 1. Fetal lung gene array data for *AGER* expression across Pseudoglandular and Canalicular stages.**

| Probe ID | AveExpr | t | *P* value | Adj *P* value | Beta-coeffient |
|----------|---------|------|-----------|---------------|----------------|
| 217046_s_at | 5.6868 | 2.985 | 0.0051 | 0.0251 | 0.0066 |
| 210081_at | 6.3418 | 5.0807 | 9.51E-06 | 0.0002 | 0.0151 |

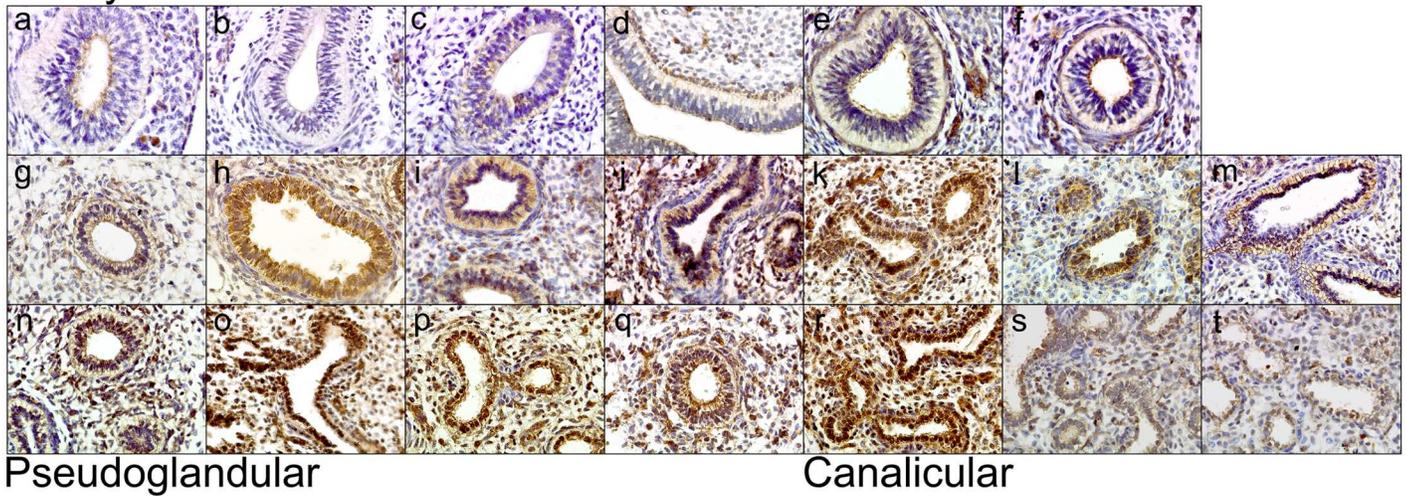doi:10.1371/journal.pone.0164041.t001

## Embryonic



**Fig 3. Immunohistochemical analysis of RAGE's expression in fetal lung.** Fetal lung tissue samples from embryonic (3–7 weeks), pseudoglandular (7–15 weeks) and canalicular (16–26 weeks) stages of development showed an increase in RAGE protein expression with increasing fetal lung age. a-t represent age ordered lung samples of 19 days (a, b), 21 days (c, d), 23 days (e, f), 9 weeks (g), 10 weeks (h,i,j), 11 weeks (k), 12 weeks (l,m,n), 13 weeks (o), 14 weeks (p), 15 weeks (q), 16 weeks (r), 17 weeks (s) and 19 weeks (t). All isotype controls were negative (data not shown). x40 magnification.

doi:10.1371/journal.pone.0164041.g003

mean change in gene expression per day during the gestational age period studied, 7–22 weeks of gestational age (data analysed as described [18]).

Having observed increasing levels of *AGER* mRNA through fetal development, levels of RAGE protein during lung development were assessed via immunohistochemistry. RAGE expression was also observed to increase during lung development across the 20 fetal lung samples; from absent to low expression in the early embryonic stages to increasing RAGE expression in the developing bronchi/bronchioles of the later embryonic stages (Fig 3a–3f). RAGE expression continued to increase across the Pseudoglandular and Canalicular stages (Fig 3g–3t).

### *AGER* mRNA levels and splice variation in airway structural cells

A total *AGER* quantitative PCR (qPCR) assay was used to determine the expression of *AGER* mRNA in human lung and in airway cell populations, specifically bronchial epithelium and airway smooth muscle cells. *AGER* was highly expressed in total lung whereas cultured human airway smooth muscle, NHBEC and BEAS2B-R1 cells exhibited lower levels of *AGER* mRNA (Fig 4). Having assessed the relative abundance of *AGER* transcripts between these cell types, we then assessed which isoforms were expressed. A large number of differentially spliced *AGER* transcripts have been previously described in lung tissue [25]. To identify *AGER* transcripts in structural cells of the airways, RNA Sequencing (RNA Seq) was used to define *AGER* splicing and isoform abundance in cultured primary human bronchial epithelial cells from three independent donors (Fig 5). Using this approach we confirmed the presence and abundance of several known *AGER* transcripts in human bronchial epithelial cells, however there was heterogeneity in expression across the donors. Interestingly, several novel variants containing exons belonging to the neighbouring gene *PCL2*, spliced together with previously annotated *AGER* exons were identified. However, while spliced transcripts were seen in all three donors, there was inconsistency in the site of splicing and abundancy of novel variants and it remains unclear whether or not these transcripts are translated (data not shown).
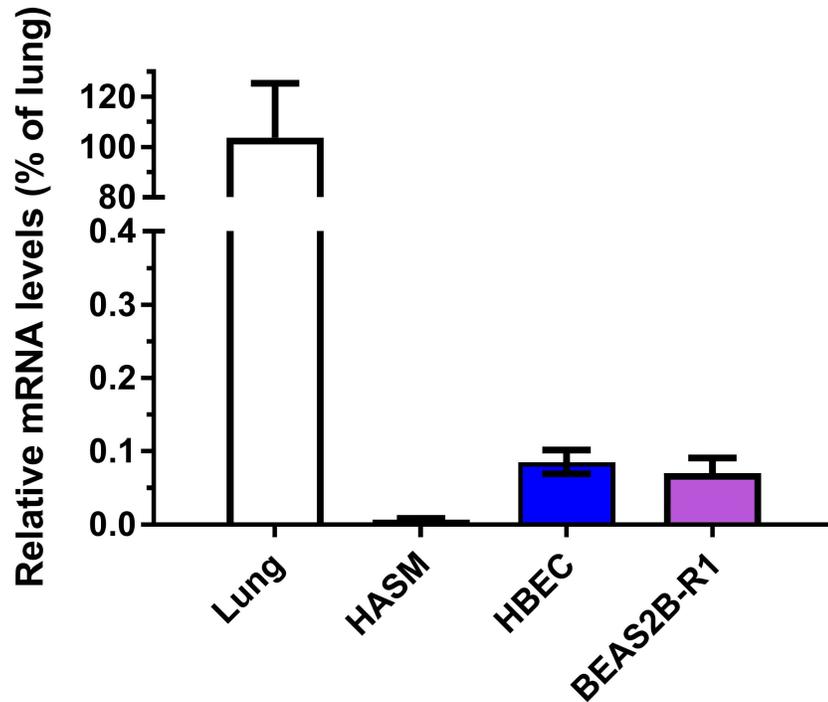
**Fig 4. *AGER* mRNA expression levels in total lung and airway cells.** Q-PCR analysis identified *AGER* mRNA was highly expressed in total lung and at lower levels in human airway smooth muscle cells (HASM) cells, human bronchial epithelial cells (HBEC) and the BEAS-2BR1 bronchial epithelial cell line, (n = 3).

| | AGER Ensembl ID | HBEC Donor 1 | HBEC Donor 2 | HBEC Donor 3 |
|---|---|---|---|---|
| | ENST00000538695 | 3.38 | 0 | 0 |
| | ENST00000488669 | 6.42 | 16.17 | 17.55 |
| | ENST00000469940 | 9.53 | 52.46 | 38.58 |
| | ENST00000484849 | 46.61 | 3.09 | 0 |
| | ENST00000473619 | 13.54 | 10.23 | 5.07 |
| | ENST00000450110 | 0.95 | 0 | 26.16 |
| | ENST00000438221 | 12.35 | 5.61 | 0 |
| | ENST00000375076* | 0 | 0 | 0 |
| | ENST0000375070 | 0 | 0 | 0 |
| | ENST00000375069 | 0 | 0 | 0 |
| | ENST00000375067 | 1.57 | 0 | 0 |
| | ENST00000375065 | 0 | 5.54 | 0 |
| | ENST00000375056 | 1.71 | 6.90 | 12.64 |
| | ENST00000375055* | 3.94 | 0 | 0 |

*ENST00000375055 soluble variant
*ENST00000375076 full length variant

**Fig 5. *AGER* isoform expression in three HBEC donors using RNA Seq.** Structure and abundance of known *AGER* isoforms in three human bronchial epithelial cell donors illustrating heterogeneity in expression levels. Percentage abundances (% FPKM) were calculated for each donor. Transcripts for full length and soluble *AGER* were identified at similar low abundances. FPKM; fragments per kilobase of transcript per million mapped reads.

**Table 2. Baseline characteristics of UK smoker population used for genetic association studies.**

| | rs2070600 genotyping (n = 1042) | Subgroup used for serum sRAGE (Gly82Gly) (n = 51) | Subgroup used for serum sRAGE (Gly82Ser) (n = 51) |
|---|---|---|---|
| Age in years (mean ± SD) | 70.0 ± 10.3 (n = 1042) | 68.2 ± 7.6 (n = 51) | 68.1 ± 7.9 (n = 51) |
| Female % (n) | 43 (n = 1042) | 31 (16 of 51) | 31 (16 of 51) |
| $FEV_1$ in litres (mean ± SD) | 1.50 ± 0.87 (n = 1036) | 1.26 ± 0.66 (n = 50) | 1.30 ± 0.72 (n = 51) |
| FEV1 percent predicted (mean ± SD) | 53.34 ± 25.47 (n = 1013) | 48.00 ± 17.79 (n = 50) | 48.68 ± 22.34 (n = 51) |
| $FEV_1$/FVC as % (mean ± SD) | 55 ± 17 (n = 1028) | 45 ± 12 (n = 49) | 48 ± 14 (n = 50) |
| Pack years (mean ± SD) | 45 ± 28 (n = 1017) | 52 ± 30 (n = 50) | 52 ± 31 (n = 49) |

doi:10.1371/journal.pone.0164041.t002

## rs2070600 is associated with lung function and sRAGE levels in the serum of a cohort of UK smokers

We analysed 1042 DNA samples from individuals who currently smoke and have a history of smoking recruited from five UK centres to assess the association of sRAGE levels with lung function by genotype. The minor allele frequency in this population for the Ser82 variant was 10%. Serum samples were analysed from 102 of these subjects who were either homozygous for the Gly82 variant or heterozygous i.e. Gly82/Ser82. Baseline characteristics of the subjects studied are shown in Table 2.

We identified association between rs2070600T and increased $FEV_1$ and $FEV_1$/FVC ratio ($P = 0.031$ and $P = 0.028$, respectively), (S1 Table). Serum from 102 subjects in the genotyping cohort was selected based on rs2070600 genotype; providing 51 heterozygotes (C/T) and 51 homozygotes (C/C) for comparison (T/T subjects not available). These groups were matched for age, sex and smoking pack years (Table 2). Serum sRAGE levels were significantly lower in individuals with the Ser82 genotype compared to individuals homozygous for Gly82 ($P < 0.0001$) showing that that the Ser82 variant leads to reduced sRAGE levels *in vivo* (Fig 6(i) and S2 Table). We did not identify a significant association between serum sRAGE levels and $FEV_1$, FVC or $FEV_1$/FVC when subjects were stratified by genotype potentially due to low numbers of subjects (S3 Table).
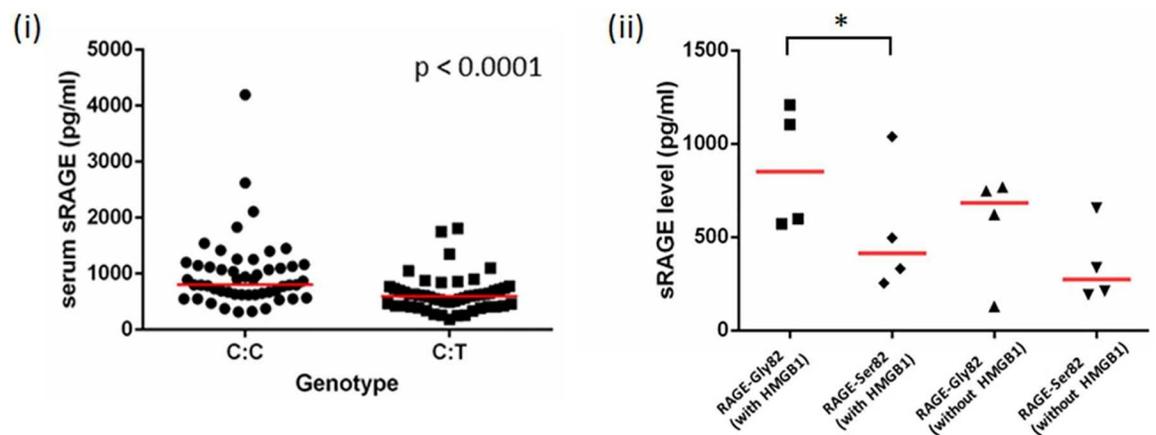


**Fig 6. The genotype effect of rs2070600.** Serum sRAGE levels are lower in individuals with the C:T genotype at rs2070600 (p<0.0001, Mann-Whitney) (6i). Cells expressing RAGE-Ser82 (T allele) stimulated with HMGB1 for 48 hours had lower levels of sRAGE than RAGE-Gly82 (C allele) carrying cells (n = 4, p = 0.036), (6ii).

doi:10.1371/journal.pone.0164041.g006

## Activation of RAGE-Ser82 leads to lower sRAGE levels compared to RAGE-Gly82 in an epithelial cell model

To investigate the functional significance of SNP rs2070600 in the context of the full length membrane bound receptor, we created a recombinant RAGE epithelial cell model, with cells engineered to express the full length RAGE coding region with either the T allele (Ser82) or the C allele (Gly82) of rs2070600. Transient transfection of allele-specific RAGE plasmids into BEAS2B-R1 (bronchial epithelial) cells resulted in a significant induction of *AGER* mRNA (mean 2500 fold change at the mRNA level) suggesting confounding from the very low endogenous expression of RAGE is unlikely to influence our findings. Stimulation with or without HMGB1 resulted in differing levels of sRAGE secretion in a genotype specific manner (Fig 6(ii)). Cells transfected with the *AGER*-Ser82 plasmid produced lower levels of sRAGE in the presence of HMGB1 compared to RAGE-Gly82 transfected cells, *P* = 0.036 (Fig 6(ii)). We did not identify a significant difference in the levels of total *AGER* mRNA in RAGE-Ser82 versus RAGE-Gly82 expressing cells (data not shown) suggesting that the effect of the T allele of rs2070600 is not due to altered transcriptional activity.

## Discussion

A number of polymorphisms within the *AGER* locus on chromosome 6p21 have shown reproducible association with the lung function measures $FEV_1$ and $FEV_1/FVC$ [1,2] and importantly COPD [5, 26]. The sentinel SNP at this locus is rs2070600, and the T allele at this SNP alters the amino acid sequence of RAGE from a Glycine at position 82 to a Serine, making this the most obvious candidate driving functional effects at this locus. In the current study we determined the expression profile of RAGE in healthy, diseased and the developing human lung demonstrating high protein expression in the alveoli of healthy adults and individuals with COPD. Importantly, we have shown for the first time that *AGER*/RAGE expression increases across developmental stages in human lung, both at the level of mRNA and protein, with high expression maintained in adulthood. We characterised *AGER* gene structure and transcript abundance in bronchial epithelial cells revealing many isoforms confirming the complexity of splice variation for *AGER* however we can not exclude that additional variation exists as we only investigated samples from three donors. Splice variants for *AGER* have previously been identified, with at least 10 coding for protein, however, only some are validated [25]. We confirmed that the rs2070600T allele (Ser82) is associated with higher values of $FEV_1$ and $FEV_1/FVC$ ratio and provide new data demonstrating that this allele is associated with lower serum levels of soluble RAGE in UK smokers in keeping with findings of two previous studies [5, 27]. Levels of sRAGE have also been associated with a diagnosis of COPD, GOLD classification and several measures of emphysema using computed tomography (CT) variables [5, 28–30]. Using an allele-specific RAGE cell model we identified that activation of the RAGE-Ser82 receptor with HMGB1 ligand leads to secretion of lower levels of soluble RAGE compared to the RAGE-Gly82 receptor demonstrating that rs2070600 is functionally relevant and could underlie the effects on serum sRAGE observed *in vivo*. This effect appears to be due to altered cell processing of RAGE, as there was no obvious effect of the variants on levels of transcription.

Our immunohistochemical data confirm and extend previously published findings that showed moderate to high expression of RAGE in the alveolar region, with low expression in the bronchial epithelium [31, 32] and higher expression in lung tissue from patients with COPD[9, 10]. We observed only weak staining in bronchial epithelium. This is consistent with the low levels of mRNA found by qPCR in NHBEC and BEAS2B-R1 cells. RAGE expression has previously been shown to increase throughout development in neonatal mouse and rat lung [33, 34]. Therefore our novel finding that RAGE expression increases across human lung

developmental stages both complement and extend the experiments in rodents. In keeping with the evidence presented here that RAGE expression increases with fetal age, Lopez Diez *et al*, showed high RAGE expression in 37 week old fetal lung [35]. For the *AGER* mRNA analyses across fetal stages it is important to note that using two probes to measure gene expression is a strength and based on the data of Hudson *et al*. [25] the 210081_at probe (exon 8) would measure 99% of adult lung transcripts and the 217046s probe (Exon 7–11) would measure 83% as this region spans alternative splicing forms. This is potentially reflected in the different intensities for these probes observed (Fig 4). A more recent study also suggested that the majority of *AGER* transcripts identified in both adult and fetal lung tissue contain exon 8 and so would be quantified by the 210081_at probe (Lopez-Diez 2013).

A limitation of the fetal immunohistochemistry was the low number of samples at each developmental stage (especially canalicular). The relationship between increasing RAGE expression in the fetal lung with increasing age may potentially be explained by cell differentiation and the evolving cellular composition within the developing lung. During the embryonic stage, early precursor lung cells are present and with increasing differentiation at the pseudoglandular stage, alveolar type II pneumocytes are present which express RAGE. Furthermore, at the canalicular stage of lung development type I pneumocytes are present alongside a larger differentiated population of type II pneumocytes expressing RAGE.

Importantly, a functional role for RAGE during lung development has recently been demonstrated in mice, with alterations in RAGE levels causing lethality and/or changes to secondary septation, hypoplasia, lung tissue loss and changes in pulmonary microvasculature [15, 36–38]. While we cannot evaluate these mechanisms in the human context our differential expression data in human fetal lung implies a role for *AGER* in mammalian lung development. Taken together, these data show that RAGE expression increases through lung development and that high expression is maintained in the adult lung suggesting a role for RAGE in lung homeostasis. This has implications for the mechanisms underlying the development of COPD as these finding potentially suggest that individuals with alterations in *AGER* homeostasis may have defective lung development and therefore never reach maximal lung function. This is at least in part supported by the association of *AGER* SNPs and lung function in children [2]. One limitation of the immunohistochemical study of fetal lung tissue was the availability of only a relatively small number of samples at each developmental stage prohibiting more quantitative or semi-quantitative analyses. A separate limitation of the study was that we expression profiled *AGER* mRNA expression in the BEAS2B-R1 immortalised cell line rather than in lung tissue or primary cells which would be optimal.

Overall, our data suggest that the Gly82Ser polymorphism alters ligand-dependent release of sRAGE in the airway. These data compliment and extend previous reports in non-airway cell systems where it has been shown that Gly82Ser can influence ligand binding and/or receptor activation [12, 39, 40]. RAGE-Ser82 receptors expressed on CHO [12] or COS7 [39] cells have a lower dissociation constant ($K_d$) for the RAGE ligand EN-RAGE or glycolaldehyde-derived AGE respectively compared to RAGE-Gly82 receptors. This reduced $K_d$ for glycolaldehyde-derived AGE for RAGE-Ser82 receptors resulted in an enhanced VEGF mRNA induction demonstrating downstream effects [39]. Subsequently, Park and colleagues demonstrated that RAGE-Ser82 was consistently glycosylated at Asn81 *in vitro* whereas RAGE-Gly82 was not [40]. This N-glycosylation was shown to be critical for the regulation of ligand (S100B) binding with potential consequences for NFKB pathway activation in HEK293 cells expressing RAGE [40]. In our epithelial cell system the RAGE-Ser82 variant was significantly associated with decreased levels of secretion of sRAGE by the cell in response to activation with the ligand HMGB1. One limitation of the cell experiments is that we did not also measure NFKB pathway activation which would have provided greater depth to our findings. However, taken

together, these data suggest that rs2070600 is likely to be the functional variant which accounts for the observed association with lung function variables as our *in vitro* model only contains this specific variant change, as opposed to other SNPs in linkage disequilibrium (LD) with rs2070600 [5].

It has been robustly demonstrated that sRAGE levels in serum or plasma of COPD patients is lower compared to control subjects and associated with GOLD stage and markers of emphysema (reviewed by Yonchuk et al 2015)[41]. This led to the concept that this lower level of sRAGE, which acts as a decoy receptor to limit inflammation may lead to greater RAGE driven inflammation in COPD. A limitation of our study is that by using samples from a general population of smokers with and without COPD we did not have data on the extent of emphysema in the individuals with COPD. Additional studies investigating the distribution of RAGE/sRAGE within the lung would help (i) clarify the source(s) of serum sRAGE (potentially alveolar epithelial cells or pulmonary capillary endothelial cells) and (ii) identify if the relative contribution of these sources is altered in COPD.

The relationship between *AGER* genetics, lung function and sRAGE levels is potentially counterintuitive as the rs2070600T allele is associated with better lung function in the general population and lower sRAGE. The molecular mechanisms underlying these observations still remain to be defined, however our novel data provide supporting evidence that rs2070600 is the key SNP driving sRAGE levels in the cell assays. The key SNP(s) driving the association with lung function and COPD diagnosis at the population level remain to be elucidated and similarly whether rs2070600 contributes to lung function and sRAGE via overlapping or distinct mechanisms is an important, unresolved question.

In summary, we provide important new information regarding the expression profile of RAGE in the human lung by identifying the isoforms present, and by assessing differential expression in human fetal lung development and in human airways from control individuals and those with COPD. The T allele of rs2070600, which codes for the RAGE-Ser82 variant, was found to be associated with increased lung function in a cohort of UK smokers and the levels of sRAGE were found to be lower in the serum of smoking individuals carrying this variant. These data provide novel insight into the role of *AGER*/RAGE in lung biology and support a functional role for rs2070600 in determining sRAGE levels, which may in turn explain the association seen between this variant and function.

## Supporting Information

**S1 Table. Linear Regression analysis of lung function measures and rs2070600 genotyping.** rs2070600T was associated with increased $FEV_1$ and $FEV_1$/FVC ratio. Minor allele frequency of tested population = 10%. Coded allele = T.
(PDF)

**S2 Table. Serum soluble RAGE levels in 102 individual stratified by rs2070600 genotype.** Serum sRAGE levels were lower in individuals with a C:T genotype for rs2070600.
(PDF)

**S3 Table. Linear regression analysis of serum sRAGE level and lung function measures FEV1, FVC and FEV1/FVC.** Serum sRAGE levels was not associated with $FEV_1$, FVC or $FEV_1$/FVC when subjects were stratified by genotype. F = statistical significance of the regression equation as a whole.
(PDF)

## Acknowledgments

## Author Contributions

## References

1. Hancock DB, Eijgelsheim M, Wilk JB, Gharib SA, Loehr LR, Marciante KD, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. Nature genetics. 2010; 42(1):45–52. Epub 2009/12/17. ng.500 [pii] doi: 10.1038/ng.500 PMID: 20010835.

2. Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, Obeidat M, et al. Genome-wide association study identifies five loci associated with lung function. Nature genetics. 2010; 42(1):36–44. Epub 2009/12/17. ng.501 [pii] doi: 10.1038/ng.501 PMID: 20010834.

3. Soler Artigas M, Wain LV, Repapi E, Obeidat M, Sayers I, Burton PR, et al. Effect of five genetic variants associated with lung function on the risk of chronic obstructive lung disease, and their joint effects on lung function. Am J Respir Crit Care Med. 2011; 184(7):786–95. Epub 2011/10/04. 184/7/786 [pii] doi: 10.1164/rccm.201102-0192OC PMID: 21965014.

4. Loth DW, Artigas MS, Gharib SA, Wain LV, Franceschini N, Koch B, et al. Genome-wide association analysis identifies six new loci associated with forced vital capacity. Nature genetics. 2014; 46(7):669–77. doi: 10.1038/ng.3011 PMID: 24929828; PubMed Central PMCID: PMC4140093.

5. Cheng DT, Kim DK, Cockayne DA, Belousov A, Bitter H, Cho MH, et al. Systemic Soluble Receptor for Advanced Glycation Endproducts Is a Biomarker of Emphysema and Associated with AGER Genetic Variants in Patients with Chronic Obstructive Pulmonary Disease. Am J Respir Crit Care Med. 2013; 188(8):948–57. Epub 2013/08/21. doi: 10.1164/rccm.201302-0247OC PMID: 23947473.

6. Yin NC, Lang XP, Wang XD, Liu W. AGER genetic polymorphisms increase risks of breast and lung cancers. Genet Mol Res. 2015; 14(4):17776–87. doi: 10.4238/2015.December.22.2 PMID: 26782423.

7.  Hori O, Brett J, Slattery T, Cao R, Zhang J, Chen JX, et al. The receptor for advanced glycation end products (RAGE) is a cellular binding site for amphoterin. Mediation of neurite outgrowth and co-expression of rage and amphoterin in the developing nervous system. The Journal of biological chemistry. 1995; 270(43):25752–61. PMID: 7592757.

8.  Sparvero LJ, Asafu-Adjei D, Kang R, Tang D, Amin N, Im J, et al. RAGE (Receptor for Advanced Glycation Endproducts), RAGE ligands, and their role in cancer and inflammation. Journal of translational medicine. 2009; 7:17. doi: 10.1186/1479-5876-7-17 PMID: 19292913; PubMed Central PMCID: PMC2666642.

9.  Wu L, Ma L, Nicholson LF, Black PN. Advanced glycation end products and its receptor (RAGE) are increased in patients with COPD. Respir Med. 2011; 105(3):329–36. Epub 2010/11/30. S0954-6111 (10)00475-0 [pii] doi: 10.1016/j.rmed.2010.11.001 PMID: 21112201.

10. Ferhani N, Letuve S, Kozhich A, Thibaudeau O, Grandsaigne M, Maret M, et al. Expression of high-mobility group box 1 and of receptor for advanced glycation end products in chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2010; 181(9):917–27. Epub 2010/02/06. 200903-0340OC [pii] doi: 10.1164/rccm.200903-0340OC PMID: 20133931.

11. Yan SD, Chen X, Fu J, Chen M, Zhu H, Roher A, et al. RAGE and amyloid-beta peptide neurotoxicity in Alzheimer's disease. Nature. 1996; 382(6593):685–91. Epub 1996/08/22. doi: 10.1038/382685a0 PMID: 8751438.

12. Hofmann MA, Drury S, Hudson BI, Gleason MR, Qu W, Lu Y, et al. RAGE and arthritis: the G82S polymorphism amplifies the inflammatory response. Genes Immun. 2002; 3(3):123–35. Epub 2002/06/19. doi: 10.1038/sj.gene.6363861 PMID: 12070776.

13. Bierhaus A, Nawroth PP. Multiple levels of regulation determine the role of the receptor for AGE (RAGE) as common soil in inflammation, immune responses and diabetes mellitus and its complications. Diabetologia. 2009; 52(11):2251–63. Epub 2009/07/29. doi: 10.1007/s00125-009-1458-9 PMID: 19636529.

14. Uhlen M, Bjorling E, Agaton C, Szigyarto CA, Amini B, Andersen E, et al. A human protein atlas for normal and cancer tissues based on antibody proteomics. Molecular & cellular proteomics: MCP. 2005; 4 (12):1920–32. doi: 10.1074/mcp.M500279-MCP200 PMID: 16127175.

15. Fineschi S, De Cunto G, Facchinetti F, Civelli M, Imbimbo BP, Carnini C, et al. Receptor for advanced glycation end products contributes to postnatal pulmonary development and adult lung maintenance program in mice. Am J Respir Cell Mol Biol. 2013; 48(2):164–71. Epub 2012/11/13. rcmb.2012-0111OC [pii] doi: 10.1165/rcmb.2012-0111OC PMID: 23144333.

16. Stogsdill MP, Stogsdill JA, Bodine BG, Fredrickson AC, Sefcik TL, Wood TT, et al. Conditional overexpression of receptors for advanced glycation end-products in the adult murine lung causes airspace enlargement and induces inflammation. Am J Respir Cell Mol Biol. 2013; 49(1):128–34. Epub 2013/03/26. rcmb.2013-0013OC [pii] doi: 10.1165/rcmb.2013-0013OC PMID: 23526218.

17. Hodge E, Nelson CP, Miller S, Billington CK, Stewart CE, Swan C, et al. HTR4 gene structure and altered expression in the developing lung. Respir Res. 2013; 14:77. Epub 2013/07/31. 1465-9921-14-77 [pii] doi: 10.1186/1465-9921-14-77 PMID: 23890215.

18. Melen E, Kho AT, Sharma S, Gaedigk R, Leeder JS, Mariani TJ, et al. Expression analysis of asthma candidate genes during human and murine lung development. Respir Res. 2011; 12:86. Epub 2011/06/28. 1465-9921-12-86 [pii] doi: 10.1186/1465-9921-12-86 PMID: 21699702.

19. Kho AT, Bhattacharya S, Tantisira KG, Carey VJ, Gaedigk R, Leeder JS, et al. Transcriptomic analysis of human lung development. Am J Respir Crit Care Med. 2010; 181(1):54–63. Epub 2009/10/10. 200907-1063OC [pii] doi: 10.1164/rccm.200907-1063OC PMID: 19815808.

20. Stewart CE, Nijmeh HS, Brightling CE, Sayers I. uPAR regulates bronchial epithelial repair in vitro and is elevated in asthmatic epithelium. Thorax. 2012; 67(6):477–87. Epub 2011/12/06. thoraxjnl-2011-200508 [pii] doi: 10.1136/thoraxjnl-2011-200508 PMID: 22139533.

21. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols. 2012; 7 (3):562–78. doi: 10.1038/nprot.2012.016 PMID: 22383036; PubMed Central PMCID: PMC3334321.

22. Stewart CE, Hall IP, Parker SG, Moffat MF, Wardlaw AJ, Connolly MJ, et al. PLAUR polymorphisms and lung function in UK smokers. BMC medical genetics. 2009; 10:112. Epub 2009/11/03. 1471-2350-10-112 [pii] doi: 10.1186/1471-2350-10-112 PMID: 19878584.

23. Portelli MA, Siedlinski M, Stewart CE, Postma DS, Nieuwenhuis MA, Vonk JM, et al. Genome-wide protein QTL mapping identifies human plasma kallikrein as a post-translational regulator of serum uPAR levels. FASEB journal: official publication of the Federation of American Societies for Experimental Biology. 2013. Epub 2013/11/20. doi: 10.1096/fj.13-240879 PMID: 24249636.

24. Reddel RR, Ke Y, Kaighn ME, Malan-Shibley L, Lechner JF, Rhim JS, et al. Human bronchial epithelial cells neoplastically transformed by v-Ki-ras: altered response to inducers of terminal squamous differentiation. Oncogene Res. 1988; 3(4):401–8. PMID: 3067190.

25. Hudson BI, Carter AM, Harja E, Kalea AZ, Arriero M, Yang H, et al. Identification, classification, and expression of RAGE gene splice variants. FASEB journal: official publication of the Federation of American Societies for Experimental Biology. 2008; 22(5):1572–80. Epub 2007/12/20. fj.07-9909com [pii] doi: 10.1096/fj.07-9909com PMID: 18089847.

26. Hobbs BD, Parker MM, Chen H, Lao T, Hardin M, Qiao D, et al. Exome Array Analysis Identifies A Common Variant in IL27 Associated with Chronic Obstructive Pulmonary Disease. Am J Respir Crit Care Med. 2016. doi: 10.1164/rccm.201510-2053OC PMID: 26771213.

27. Gaens KH, Ferreira I, van der Kallen CJ, van Greevenbroek MM, Blaak EE, Feskens EJ, et al. Association of polymorphism in the receptor for advanced glycation end products (RAGE) gene with circulating RAGE levels. J Clin Endocrinol Metab. 2009; 94(12):5174–80. Epub 2009/11/06. jc.2009-1067 [pii] doi: 10.1210/jc.2009-1067 PMID: 19890027.

28. Miniati M, Monti S, Basta G, Cocci F, Fornai E, Bottai M. Soluble receptor for advanced glycation end products in COPD: relationship with emphysema and chronic cor pulmonale: a case-control study. Respir Res. 2011; 12:37. Epub 2011/04/01. 1465-9921-12-37 [pii] doi: 10.1186/1465-9921-12-37 PMID: 21450080.

29. Gopal P, Rutten EP, Dentener MA, Wouters EF, Reynaert NL. Decreased plasma sRAGE levels in COPD: influence of oxygen therapy. Eur J Clin Invest. 2012; 42(8):807–14. Epub 2012/02/01. doi: 10.1111/j.1365-2362.2012.02646.x PMID: 22288943.

30. Cockayne DA, Cheng DT, Waschki B, Sridhar S, Ravindran P, Hilton H, et al. Systemic biomarkers of neutrophilic inflammation, tissue injury and repair in COPD patients with differing levels of disease severity. PLoS One. 2012; 7(6):e38629. Epub 2012/06/16. doi: 10.1371/journal.pone.0038629 PONE-D-11-24431 [pii]. PMID: 22701684.

31. Katsuoka F, Kawakami Y, Arai T, Imuta H, Fujiwara M, Kanma H, et al. Type II alveolar epithelial cells in lung express receptor for advanced glycation end products (RAGE) gene. Biochem Biophys Res Commun. 1997; 238(2):512–6. Epub 1997/09/23. S0006-291X(97)97263-6 [pii] doi: 10.1006/bbrc.1997.7263 PMID: 9299542.

32. Shirasawa M, Fujiwara N, Hirabayashi S, Ohno H, Iida J, Makita K, et al. Receptor for advanced glycation end-products is a marker of type I lung alveolar cells. Genes Cells. 2004; 9(2):165–74. Epub 2004/03/11. 712 [pii]. PMID: 15009093.

33. Reynolds PR, Kasteler SD, Cosio MG, Sturrock A, Huecksteadt T, Hoidal JR. RAGE: developmental expression and positive feedback regulation by Egr-1 during cigarette smoke exposure in pulmonary epithelial cells. Am J Physiol Lung Cell Mol Physiol. 2008; 294(6):L1094–101. Epub 2008/04/09. 00318.2007 [pii] doi: 10.1152/ajplung.00318.2007 PMID: 18390831.

34. Lizotte PP, Hanford LE, Enghild JJ, Nozik-Grayck E, Giles BL, Oury TD. Developmental expression of the receptor for advanced glycation end-products (RAGE) and its response to hyperoxia in the neonatal rat lung. BMC Dev Biol. 2007; 7:15. Epub 2007/03/09. 1471-213X-7-15 [pii] doi: 10.1186/1471-213X-7-15 PMID: 17343756.

35. Lopez-Diez R, Rastrojo A, Villate O, Aguado B. Complex tissue-specific patterns and distribution of multiple RAGE splice variants in different mammals. Genome Biol Evol. 2013; 5(12):2420–35. doi: 10.1093/gbe/evt188 PMID: 24273313; PubMed Central PMCID: PMC3879976.

36. Winden DR, Ferguson NT, Bukey BR, Geyer AJ, Wright AJ, Jergensen ZR, et al. Conditional overexpression of RAGE by embryonic alveolar epithelium compromises the respiratory membrane and impairs endothelial cell differentiation. Respir Res. 2013; 14(1):108. Epub 2013/10/19. 1465-9921-14-108 [pii] doi: 10.1186/1465-9921-14-108 PMID: 24134692.

37. Stogsdill JA, Stogsdill MP, Porter JL, Hancock JM, Robinson AB, Reynolds PR. Embryonic overexpression of receptors for advanced glycation end-products by alveolar epithelium induces an imbalance between proliferation and apoptosis. Am J Respir Cell Mol Biol. 2012; 47(1):60–6. Epub 2012/02/22. rcmb.2011-0385OC [pii] doi: 10.1165/rcmb.2011-0385OC PMID: 22343220.

38. Reynolds PR, Stogsdill JA, Stogsdill MP, Heimann NB. Up-regulation of receptors for advanced glycation end-products by alveolar epithelium influences cytodifferentiation and causes severe lung hypoplasia. Am J Respir Cell Mol Biol. 2011; 45(6):1195–202. Epub 2011/06/21. 2011-0170OC [pii] doi: 10.1165/rcmb.2011-0170OC PMID: 21685154.

39. Osawa M, Yamamoto Y, Munesue S, Murakami N, Sakurai S, Watanabe T, et al. De-N-glycosylation or G82S mutation of RAGE sensitizes its interaction with advanced glycation endproducts. Biochimica et biophysica acta. 2007; 1770(10):1468–74. doi: 10.1016/j.bbagen.2007.07.003 PMID: 17714874.

40. Park SJ, Kleffmann T, Hessian PA. The G82S polymorphism promotes glycosylation of the receptor for advanced glycation end products (RAGE) at asparagine 81: comparison of wild-type rage with the

G82S polymorphic variant. The Journal of biological chemistry. 2011; 286(24):21384–92. Epub 2011/ 04/23. M111.241281 [pii] doi: 10.1074/jbc.M111.241281 PMID: 21511948.

41. Yonchuk JG, Silverman EK, Bowler RP, Agusti A, Lomas DA, Miller BE, et al. Circulating soluble receptor for advanced glycation end products (sRAGE) as a biomarker of emphysema and the RAGE axis in the lung. Am J Respir Crit Care Med. 2015; 192(7):785–92. doi: 10.1164/rccm.201501-0137PP PMID: 26132989.