ERJ open research
THE BEST IN OPEN ACCESS BASIC, TRANSLATIONAL & CLINICAL RESPIRATORY RESEARCH

Early View

Original research article

Evidence for shared genetic risk factors between lymphangioleiomyomatosis and pulmonary function

Xavier Farré, Roderic Espín, Alexandra Baiges, Eline Blommaert, Wonji Kim, Krinio Giannikou, Carmen Herranz, Antonio Román, Berta Sáez, Álvaro Casanova, Julio Ancochea, Claudia Valenzuela, Piedad Ussetti, Rosalía Laporta, José A. Rodríguez-Portal, Coline H.M. van Moorsel, Joanne J. van der Vis, Marian J.R. Quanjel, Mireia Tena-Garitaonaindia, Fermín Sánchez de Medina, Francesca Mateo, María Molina-Molina, Sungho Won, David J. Kwiatkowski, Rafael de Cid, Miquel Angel Pujana

Please cite this article as: Farré X, Espín R, Baiges A, *et al*. Evidence for shared genetic risk factors between lymphangioleiomyomatosis and pulmonary function. *ERJ Open Res* 2021; in press (https://doi.org/10.1183/23120541.00375-2021).

Evidence for shared genetic risk factors between lymphangioleiomyomatosis and pulmonary function

Xavier Farré[1,17], Roderic Espín[2,17], Alexandra Baiges[2,17], Eline Blommaert[2], Wonji Kim[3], Krinio Giannikou[4], Carmen Herranz[2], Antonio Román[5], Berta Sáez[5], Álvaro Casanova[6], Julio Ancochea[7], Claudia Valenzuela[7], Piedad Ussetti[8], Rosalía Laporta[8], José A. Rodríguez-Portal[9,10], Coline H.M. van Moorsel[11], Joanne J. van der Vis[11], Marian J.R. Quanjel[11], Mireia Tena-Garitaonaindia[12,13], Fermín Sánchez de Medina[13,14], Francesca Mateo[2], María Molina-Molina[10,15], Sungho Won[16], David J. Kwiatkowski[4,18], Rafael de Cid[1,18] and Miquel Angel Pujana[2,18]

1 Genomes for Life – GCAT Lab Group, Institut Germans Trias i Pujol (IGTP), Badalona, Catalonia, Spain.
2 ProCURE, Catalan Institute of Oncology, Oncobell, Bellvitge Institute for Biomedical Research (IDIBELL), L'Hospitalet del Llobregat, Barcelona, Catalonia, Spain.
3 Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA.
4 Cancer Genetics Laboratory, Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA.
5 Lung Transplant Unit, Pneumology Service, Lymphangioleiomyomatosis Clinic, Vall d'Hebron University Hospital, Vall d'Hebron Research Institute (VHIR), Barcelona, Spain.
6 Pneumology Service, University Hospital of Henares, University Francisco de Vitoria, Coslada, Madrid, Spain.
7 Pneumology Service, La Princesa Research Institute, University Hospital La Princesa, Madrid, Spain.
8 Pneumology Service, University Hospital Clinica Puerta del Hierro, Majadahonda, Madrid, Spain.
9 Medical-Surgical Unit of Respiratory Diseases, University Hospital Virgen del Rocío, Institute of Biomedicine of Seville (IBiS), Seville, Spain.
10 Biomedical Research Network Centre in Respiratory Diseases (CIBERES), Instituto de Salud Carlos III, Madrid, Spain.
11 Interstitial Lung Disease (ILD) Center of Excellence, St. Antonius Hospital, Nieuwegein, The Netherlands.
12 Department of Biochemistry and Molecular Biology II, School of Pharmacy, Instituto de Investigación Biosanitaria ibs.GRANADA, University of Granada, Granada, Spain.
13 Biomedical Research Network Centre in Hepatic and Digestive Diseases (CIBERehd), Instituto de Salud Carlos III, Madrid, Spain.
14 Department of Pharmacology, School of Pharmacy, Instituto de Investigación Biosanitaria ibs.GRANADA, University of Granada, Granada, Spain.
15 Interstitial Lung Disease Unit, Department of Respiratory Medicine, University Hospital of Bellvitge, IDIBELL, L'Hospitalet del Llobregat, Barcelona, Spain.
16 Department of Public Health Sciences, Interdisciplinary Program of Bioinformatics, Institute of Health and Environment, Seoul National University, Seoul, South Korea.
17 These two authors contributed equally as first authors.
18 Joint senior authors.

Correspondence:
Miquel Angel Pujana, ProCURE, Catalan Institute of Oncology, IDIBELL, Gran via 199, L'Hospitalet del Llobregat, Barcelona 08908, Catalonia, Spain. Ph. +34-932607952; E-mail: mapujana@iconcologia.net
Rafael de Cid, Genomes for Life – GCAT Lab Group, Institut Germans Trias i Pujol (IGTP), Carretera Can Ruti, Camí de les Escoles s/n, Badalona 08916, Catalonia, Spain. Ph. +34-935543050; E-mail: rdecid@igtp.cat

Conflict of interest: The authors declare no conflict of interests.

Abstract.

Introduction: Lymphangioleiomyomatosis (LAM) is a rare low-grade metastasizing disease characterized by cystic lung destruction. The genetic basis of LAM remains incompletely determined, and the disease cell-of-origin is uncertain. We analysed the possibility of a shared genetic basis between LAM and cancer, and LAM and pulmonary function. **Methods:** The results of genome-wide association studies (GWASs) of LAM, 17 cancer types, and spirometry measures (forced expiratory volume in 1-second (FEV_1), forced vital capacity (FVC), FEV_1/FVC ratio, and peak expiratory flow (PEF)) were analysed for genetic correlations, shared genetic variants, and causality. Genomic and transcriptomic data were examined, and immunodetection assays were performed to evaluate pleiotropic genes. **Results:** There were no significant overall genetic correlations between LAM and cancer, but LAM correlated negatively with FVC and PEF, and a trend in the same direction was observed for FEV_1. Twenty-two shared genetic variants were uncovered between LAM and pulmonary function, while seven shared variants were identified between LAM and cancer. The LAM-pulmonary function shared genetics identified four pleiotropic genes previously recognized in LAM single-cell transcriptomes: *ADAM12*, *BNC2*, *NR2F2*, and *SP5*. We had previously associated *NR2F2* variants with LAM, and we identified its functional partner NR3C1 as another pleotropic factor. NR3C1 expression was confirmed in LAM lung lesions. Another candidate pleiotropic factor, CNTN2, was found more abundant in plasma of LAM patients than that of healthy women. **Conclusions:** This study suggests the existence of a common genetic aetiology between LAM and pulmonary function.

Introduction

Lymphangioleiomyomatosis (LAM) is a rare low-grade progressive neoplasm that affects women almost exclusively and is characterized by cystic lung destruction, which can lead to respiratory failure in severe cases (1, 2). In addition to cystic lung disease, LAM is also strongly associated with renal angiomyolipomas (AML) and lymphatic alterations, for which reason it may be properly considered a systemic disorder (1, 2). LAM lung lesions are characterized by low-grade proliferation of 'LAM cells' that have both smooth muscle cell-like features, and microphthalmia-associated transcription factor (MITF)-driven gene expression; and by cyst formation, which is likely driven by expression of proteases and cathepsins (3–5). The tissue of origin of LAM cells is uncertain (6–8).

LAM can occur sporadically (S-LAM) or in the presence of Tuberous Sclerosis Complex (TSC-LAM), an autosomal-dominant multisystem disorder caused by heterozygous germline or mosaic loss-of-function mutations in the tumour suppressor genes *TSC1* and *TSC2* (9, 10). In sporadic LAM (S-LAM), somatic inactivation of *TSC2*, or much less commonly *TSC1*, in an unknown cell type(s) appears to be sufficient for disease development (11). Germline or somatic mutations in *TSC1/TSC2* lead to hyperactivation of the mechanistic target of rapamycin complex 1 (mTORC1), and mTOR allosteric inhibitors (rapamycin/sirolimus and its derivates, rapalogs) are the main therapeutic approach for LAM (12). Due to its central role in metabolism, mTORC1 activity is abnormally enhanced in many cancer types, and generally linked to stem cell-like features, which are also present in LAM cells (13–16). Indeed, LAM shows several fundamental hallmarks of cancer, including continued cell proliferation and resistance to cell death, expression of factors promoting tissue invasion and metastasis, and immune evasion (6, 17–19).

A recent genome wide association study (GWAS) we performed identified several common genetic variants on chromosome 15q26.2 for which an allele was associated with risk of S-LAM (20). Although not certain, these alleles appear to affect LAM development through effects on the nuclear receptor subfamily 2 group F member 2 (*NR2F2*) gene, also known as chicken ovalbumin upstream promoter transcription factor II (COUP-TFII). This transcription factor is widely expressed during embryogenesis, and has a role in various endocrine conditions and cancers (21, 22). Here we sought to explore in further detail the genetic basis of LAM risk, considering the hypothesis that there is a shared genetic basis between LAM and cancer and/or pulmonary function.

Materials and methods

GWAS data

The GWAS summary statistics of LAM (20), 17 cancer types (23) and pulmonary function tests (including of FEV_1, FVC, FEV_1/FVC ratio, and PEF) (24) were obtained from the corresponding data sources (**Table S1** provides details of the GWASs, including sample sizes and relevant references). This study did not require individual data. The original LAM GWAS included 426 S-LAM subjects that were analysed in comparison with 852 females from the COPDGene study in a matched case-control design (20).

Imputation and data preprocessing

The GWAS summary statistics for LAM were imputed using the SSimp software (25, 26) and European individuals from the 1000 Genomes Project phase 3 reference panel, filtering out single nucleotide polymorphisms (SNPs) with a minor allele frequency (MAF) ≤ 0.01. The imputation increased the number of SNPs from 681,894 to 9,325,933, but those with poor imputation quality (r^2 < 0.3) were removed, providing a set of 7,809,072 SNPs for subsequent analyses. For each of the variant-summary statistics, standard quality controls were applied: removal of SNPs without reference identifier (rs ID), duplicated, poorly imputed (info score < 0.9), with MAF ≤ 0.01, with strand ambiguous alleles and/or with sample size five standard deviations away from the mean. In addition, SNPs in the extended major histocompatibility complex (hg 19 coordinates, chr6: 25,119,106–33,854,733) and 8p23.1 region (chr8: 7,200,000–12,500,000) were excluded given that these regions have complex linkage disequilibrium (LD) that could bias pleiotropy analyses (27).

Shared genetic architecture

The estimation of heritability of all phenotypes and computation of genetic correlation between GWASs of LAM and each of the other selected diseases or traits were performed using the high-definition likelihood (HDL) inference method (28). Genetic correlations estimated using HDL fully account for LD across the genome, increasing the power to detect connections between complex traits using precomputed LD matrices for 335,265 Genomic British individuals in the UK Biobank and HapMap3 SNPs (28). For detection of shared genetic risk factors, we used the pleiotropy-informed conditional false discovery rate approach (29) applying pleioFDR software (https://github.com/precimed/pleiofdr/) and computing conjunctional false discovery rate (conjFDR) statistics. The conjFDR is given by the maximum between the conditional FDRs (condFDR) for two given conditions; the condFDR method was shown to improve statistical power relative to the conventional approach of using *P* value thresholds for detection of shared genetics, and was not affected by the direction of the allele effects (30, 31). A pairwise analysis was performed between LAM GWAS and each other condition. In order to make the results comparable, we analysed a common set of 5,684,891 SNPs present in all summary statistics. Shared genetic variants were defined by conjFDR < 0.05. We then performed LD clumping to define independent significant SNPs (PLINK software, p1 = 0.05; p2 = 1, LD threshold r^2 = 0.6, and physical distance threshold for clumping 1,000 kilobases (kb)) and lead SNPs (PLINK software, p1 = 0.05, p2 = 1, r^2 = 0.1, and distance 1,000 kb). Genomic risk loci were found by merging lead SNPs if they were closer than 250 kb. Candidate SNPs were mapped to independent significant SNPs using this clumping strategy. Stratified Q-Q plots were obtained using pleioFDR to visualize shared genetic architecture. In these representations, the *P* values of the primary phenotype were plotted against the null distribution. In the same plots, we represented subsets of SNPs of the primary phenotype conditioned by the significance

of their association with the secondary phenotype; *P* value of the secondary phenotype < 0.1, 0.01, and 0.001. A leftwards deflection of the lines implies the presence of shared genetic architecture.

Mendelian randomization

A two-sample Mendelian randomization (MR) approach was applied to evaluate causality using GWAS summary statistics of cancer risk and spirometry measures as exposure or outcome, and LAM as the outcome or exposure, respectively. A valid instrumental variable should fulfil three core assumptions: the variant is associated with the exposure; the variant is independent of all confounders of the exposure-outcome association; and the variant is only associated with the outcome through its effect on the exposure (32). Independent genetic variants (LD r^2 > 0.001; genome-wide significance, $P < 5 \times 10^{-8}$, for spirometry measures; and at a suggestive level, $P < 1 \times 10^{-5}$ for cancer and LAM risk) were used as instrumental variables. The three-step MR-Egger method (33) was applied: 1) a test for directional pleiotropy; 2) a test for a causal effect; and 3) estimation of the causal effect. To assess the robustness of the results, causal estimates and *P* values were compared using random-effects inverse-variance-weighted (IVW) and robust adjusted profile score (MR-RAPS) (34) methods. Heterogeneity was calculated from Cochran's Q value (35). The MR Steiger directionality test, which compares the variance explained by the SNPs in the exposure and outcome, was applied to elucidate the direction of causality (36). This test estimates directionality leveraging the fact that in most settings the genetic variants will explain more variance of the trait located upstream in the causal chain. Informative scatter and forest plots were generated to examine the results. The analyses were performed using the TwoSampleMR R package (37).

Gene candidate prioritization

Information of positional gene candidates (up to 1 Megabase from the given variant) was integrated with data of expression quantitative trait loci (eQTL) in *cis* identified in non-diseased human tissue (Genotype-Tissue Expression (GTEx) project (38)), and with data of chromatin interactions identified in lung tissue as potential cis-regulatory elements (39). The GTEx evidence was obtained from pan-tissue analyses as well as from tissue expected to be biologically related to LAM disease: blood vessel, lung, kidney and uterus. Information for physical protein interactions and complex membership was taken from the Human Reference Protein Interactome (40) and Biogrid (41) databases. For the evaluation of publicly available single-cell RNA-sequencing (RNA-seq) profiles from LAM1-4 diseased lungs (Gene Expression Omnibus reference GSE135851), the samples were pre-processed and analysed as originally described (8), using the Seurat R package (42). LAM cells (excluding lung mesenchymal cells) were identified using the LAM^{core} gene expression signature originally described by the authors as the reference (8). RNA-seq data of kidney AMLs corresponded to 14 published cases from two studies (n = 5 (20); and n = 9 (43)) and 14 unpublished cases (internal cohort, manuscript in preparation). RNA-seq data of other cancer types corresponded to The Cancer Genome Atlas (TCGA) (44) and were downloaded from the cBioPortal for Cancer Genomics (45). The single-cell RNA-seq data of the Human Lung Cell Atlas (46) was downloaded from the Synapse portal (https://www.synapse.org/#!Synapse:syn21041850/wiki/600865). For all candidate pleiotropic genes, and for each lung cell type, a percentage of cells with a given gene expression was calculated using the ComplexHeatmap R package with standard parameters (47).

LAM lung samples, immunohistochemistry

LAM patients were recruited and lung tissue samples collected by the participating centres (International LAM Clinic, University Hospital Vall d'Hebron; University Hospital La Princesa; University Hospital Clínica Puerta del Hierro; University Hospital Clínic Barcelona; University Hospital Virgen del Rocío; and University Hospital of Bellvitge), and with the support of the Spanish LAM Association (AELAM). The patients provided written informed consent and the study were approved by the Research Ethics Committee of the IDIBELL. The immunohistochemical assays were performed on serial paraffin sections using an EnVision kit (Dako) and antigen retrieved with citrate pH 6.0 buffer. Endogenous peroxidase was blocked by pre-incubation in a solution of 3% H_2O_2, performed in 1x phosphate-buffered saline with 10% goat serum. Slides were incubated overnight at 4ºC with primary antibodies (anti-NR3C1, dilution 1:100, D6H2L, Cell Signalling Technology; and anti-NTN4, dilution 1:30, HPA049832, Sigma-Aldrich) in blocking solution. Secondary peroxidase-conjugated antibody (Envision+ system-HRP, Dako) was used. Sections were counterstained with haematoxylin and examined under a Nikon Eclipse 80i microscope. For immunohistofluorescence, the slides were incubated with a mixture of the two primary antibodies (anti-αSMA, dilution 1:1000, A2547, Sigma-Aldrich; and anti-NR3C1). The secondary antibodies used were goat anti-mouse Alexa-546 and goat anti-rabbit Alexa-488 (dilution 1:100; Thermo Fisher). Sudan black staining was performed to avoid paraffin autofluorescence. The sections were mounted with coverslips in Vectashield containing DAPI and visualized under a Nikon 80i epifluorescence microscope equipped with a DS-Ri1 camera.

Plasma samples and ELISA

LAM blood samples were collected and immediately processed during the 2017 and 2018 annual AELAM patient conferences, so the time between undertaking pulmonary function tests and sample acquisition varied, making it impossible to assess the former relative to the biomarkers. The data collected consisted of age at diagnosis, age at sample extraction, diagnosis of AML, chylothorax, pneumothorax, TSC, and therapy used at the time of sample extraction. All patients provided written informed consent and the study was approved by the ethics committees of IDIBELL and the Instituto de Investigación Sanitaria La Princesa, Hospital de Henares. Control samples were obtained from healthy premenopausal women from a similar age distribution to that of the LAM patients. Lung disease-related patients corresponded to cases diagnosed with emphysema (excluding chronic obstructive pulmonary disease; University Hospital of Bellvitge, IDIBELL), Langerhans cell histiocytosis, Sjögren syndrome, and systemic lupus erythematosus. Blood samples of the former conditions were collected by the ILD Centre of Excellence, St. Antonius Hospital Biobank (Nieuwegein, The Netherlands). This study was approved by the institutional ethics committee (reference R05-08A) and all participants provided written informed consent. CNTN2 was quantified in blood plasma using the Human Contactin-2/TAG1 DuoSet enzyme-linked immunosorbent assay (ELISA; DY1714-05, R&D Systems) following the manufacturer's instructions. VEGF-D levels were measured using a commercially available ELISA kit according to the manufacturer's protocol (DVED00, R&D Systems).

Results

Identification of shared genetics of LAM and cancers

As a low-grade neoplasm, LAM has a biological similarity to cancer (6, 17–19). This connection might be extended to disease susceptibility (48, 49) and could indicate the existence of a shared genetic basis of the two conditions. To assess this hypothesis, we analysed summary statistic results from the original S-LAM GWAS (20) and from studies of 17 cancer types, including those of breast, colon, glioma, kidney, lung, neuroblastoma, ovary, pancreas, prostate, stomach, and skin (**Table S1**). After SNP imputation, data preprocessing and quality control analyses, the summary statistics of 5,684,891 SNPs were evaluated between LAM and cancers. The overall pairwise genetic correlation between LAM and cancer risk did not reach significance for any cancer; the strongest correlation was negative and for cervical cancer risk (r = -0.32, P = 0.083; **Table S2**). In addition, when conditioning LAM on any cancer type, there were no substantial signs of shared genetic architecture in the stratified Q-Q plots (**Figure S1**). However, for some cancer analyses, there was not enough statistical power to estimate genetic correlations accurately (**Table S2**).

Despite the non-significant overall LAM-cancer genetic correlation, seven shared genetic associations were identified with conjFDR < 0.05, including variants between LAM and gastric, kidney or prostate cancer risk (**Table 1**). These variants were mapped to six genomic loci, of which three were linked to gastric cancer, two to kidney cancer, and one to prostate cancer risk. Four of the associated loci (including gastric, kidney and prostate cancer) were predicted to function as agonists; that is, the corresponding minor alleles showed the same direction of effect for LAM and cancer risk; in turn, two were predicted to act as antagonists (including gastric and kidney cancer; **Table 1**).

The chromosome 6q24 rs3861451 variant shared between LAM and gastric cancer risk was relatively linked (Caucasian $D´$ = 0.74 and r^2 = 0.52) with a previously noted

SNP association for this cancer type, rs618688 (23). Next, we inspected the seven identified LAM-cancer shared variants (**Table 1**) relative to the GWAS catalogue for human traits and diseases (50, 51). This examination identified the chromosome 2q31 variant rs4668267 that connects LAM and prostate cancer risk as a genome-wide association signal of earlobe attachment (52).

Table 1. Shared genetic variants between LAM and cancer risk (ordered by cancer type).

Cancer type	SNP	Chr	Position (bp, hg19)	Reference allele	Alternative allele (MAF)	Cancer *P*	LAM *P*	Cancer *Z* score	LAM *Z* score	Pleiotropy	conjFDR	Locus
Gastric	rs3861451	6	148,320,047	T	C (0.33)	1.6×10^{-4}	1.2×10^{-5}	3.78	-4.38	Antagonist	0.022	Intergenic
	rs9528802	13	65,235,556	T	C (0.29)	1.8×10^{-4}	2.4×10^{-4}	3.74	3.68	Agonist	0.035	Intergenic
	rs10901587	10	128,001,759	C	T (0.23)	2.5×10^{-4}	2.5×10^{-4}	3.67	3.66	Agonist	0.037	*ADAM12* intronic
Kidney	rs4512050	4	158,597,491	A	G (0.13)	2.9×10^{-4}	1.9×10^{-4}	3.62	3.74	Agonist	0.044	Intergenic
	rs17036640		158,605,184	A	G (0.13)	2.5×10^{-4}	1.7×10^{-4}	3.67	3.76	Agonist	0.037	Intergenic
	rs2146084	9	11,2505,617	G	A (0.12)	2.6×10^{-5}	2.7×10^{-4}	-4.21	3.65	Antagonist	0.047	*PALM2-AKAP2* intronic
Prostate	rs4668267	2	171,360,896	T	C (0.30)	7.9×10^{-5}	8.7×10^{-5}	3.95	3.93	Agonist	0.030	*MYO3B* intronic

Identification of shared genetics of LAM and pulmonary function

In parallel with cancer, we analysed shared genetic factors between LAM and pulmonary function, determined by spirometry measures of FEV_1, FVC, FEV_1/FVC ratio, and PEF in the general population using summary statistic GWAS results from the UK Biobank and SpiroMeta consortium (24). The overall pairwise genetic correlations between LAM and FVC, and PEF, were nominally significant and negative: r = -0.074, 95% confidence interval (CI) -0.04 — -0.10, P = 0.013; and r = -0.10, 95% CI -0.05 — -0.15, P = 0.029, respectively (**Table S2**). The LAM correlation with FEV_1 was also negative, but did not reach significance: r = -0.061, P = 0.097. By conditioning LAM on spirometry measures, the stratified Q-Q plots showed a leftwards deflection with increasingly strong associations with the depicted measures, such as PEF (**Figure 1**; Q-Q plots for the rest of measures are shown in **Figure S2**).

Twenty-two variants in 11 loci were common to LAM and pulmonary function: 11 shared with FEV_1, 7 with FVC, 4 with FEV_1/FVC ratio and 6 with PEF (**Table 2**). Six of the identified variants mapped to chromosome 15q26.2, relatively close (≤ 35 kilobases (kb)) to the primary S-LAM associations targeting *NR2F2* (20); the six identified variants were in relative linkage disequilibrium (Caucasian $D' > 0.90$, $r^2 > 0.25$) with the original genome-wide significant associations detected by rs2006950 and rs4544201 (20). Inspection of the UK Biobank and SpiroMeta results showed nominal associations between these two original SNPs and pulmonary function measures, with P values between 0.05 and 0.0002. Interestingly, all six variants identified in 15q26.2 had opposite effects in pulmonary function compared with LAM risk: that is, their minor alleles were associated with relatively superior pulmonary function, but with lower LAM risk (**Table 2**).

There was no overlap between the shared variants influencing LAM and cancer, and LAM and pulmonary function. We further inspected the 22 LAM-pulmonary function shared variants relative to cancer risk associations (50, 51) in the corresponding

genomic regions. The chromosome 5q31.3 rs7701443 variant is located at 37 kb (Caucasian *D'* = 0.90, r^2 = 0.15) of a breast cancer association identified by the Breast Cancer Association Consortium, rs2963155 (53), and the chromosome 9p22.3 variant rs4961722 is located at < 1 kb (Caucasian *D'* = 1.0, r^2 = 0.03) from a skin cancer association, rs10962474 (54). In addition, the chromosome 8q23.3 region comprising five LAM-pulmonary function shared variants (**Table 2**) was relatively strongly linked (Caucasian LD *D'* = 0.25-1; maximum distance between variants ≤ 45 kb) to association signals of colorectal, gastric, and rectal cancer risk: rs6469654, rs6469656, rs16892766, rs76316943, rs117079142, and rs200235517 (55).

Next, we inspected the identified LAM-pulmonary function loci for associations with lung-related traits. The LAM-FEV_1 chromosome 2p21 rs13410076 variant is located at 230 kb from a rare variant (Caucasian MAF < 0.001) previously associated at the genome-wide level with oxygenated haemoglobin levels in Tibetan women living at high altitude, rs372272284, which may influence the expression of endothelial PAS domain containing protein 1 (*EPAS1/HIF2A*) gene (56). Evaluation of other traits identified the LAM-FEV_1/FVC chromosome 1q32.1 rs16937 and rs11240341 variants as genome-wide association signals for mean platelet volume (57) and schizophrenia (58).

Table 2. Shared genetic variants between LAM and pulmonary function (ordered by chromosome).

Trait	SNP	Chr	Position (bp, hg19)	Reference allele	Alternative Allele (MAF)	Trait *P*	LAM *P*	Trait *Z* score	LAM *Z* score	Pleiotropy	conjFDR	Gene locus
FEV_1/FVC	rs11240341	1	205,015,284	C	T (0.34)	7.7×10^{-6}	7.5×10^{-6}	-4.47	4.48	Antagonist	0.007	*CNTN2* intronic
FEV_1/FVC	rs16937		205,035,455	G	A (0.32)	1.8×10^{-5}	1.9×10^{-5}	-4.29	4.28	Antagonist	0.017	
FEV_1	rs13410076	2	46,815,961	T	C (0.09)	3.4×10^{-4}	3.9×10^{-5}	-3.58	-4.11	Agonist	0.022	*PIGF* intronic
PEF	rs835069	5	14,984,116	C	T (0.15)	3.9×10^{-4}	1.1×10^{-4}	3.54	-3.86	Antagonist	0.044	Intergenic
FEV_1	rs7701443		142,792,650	A	G (0.39)	3.6×10^{-4}	1.5×10^{-5}	3.57	4.32	Agonist	0.023	*NR3C1* intronic
FVC						2.1×10^{-7}	1.5×10^{-5}	5.19	4.32	Agonist	0.011	
FVC	rs2537572	7	17,771,780	A	G (0.41)	9.5×10^{-7}	7.5×10^{-5}	-4.90	3.96	Antagonist	0.046	Intergenic
FEV_1						4.4×10^{-4}	2.8×10^{-7}	3.51	-5.14	Antagonist	0.026	
FEV1	rs9648555		46,656,869	A	G (0.47)	7.7×10^{-7}	7.5×10^{-5}	-4.94	3.96	Antagonist	0.033	Intergenic
FVC	rs2511654	8	117,623,003	T	C (0.30)	1.7×10^{-4}	2.7×10^{-5}	-3.76	4.19	Antagonist	0.018	Intergenic
FEV_1	rs12542425		117,642,345	C	T (0.37)	2.2×10^{-4}	1.4×10^{-6}	-3.69	4.82	Antagonist	0.015	Intergenic
FVC						1.7×10^{-4}	1.4×10^{-6}	-3.77	4.82	Antagonist	0.014	
FVC	rs7839361		117,645,687	T	C (0.17)	6.4×10^{-5}	7.4×10^{-6}	-3.99	4.48	Antagonist	0.006	Intergenic
FEV_1	rs17663673		117,650,006	T	C (0.31)	2.1×10^{-4}	5.0×10^{-5}	-3.71	4.05	Antagonist	0.023	Intergenic
FVC						3.1×10^{-4}	5.0×10^{-5}	-3.61	4.05	Antagonist	0.032	
FEV_1	rs34672734		117,805,281	A	G (0.21)	3.6×10^{-4}	3.7×10^{-5}	-3.57	4.13	Antagonist	0.023	Intergenic
FVC						1.4×10^{-4}	3.7×10^{-5}	-3.82	4.13	Antagonist	0.025	
FEV_1/FVC	rs4961722	9	16,529,174	T	C (0.37)	1.8×10^{-4}	3.4×10^{-5}	3.75	4.15	Agonist	0.030	*BNC2* intronic
FEV_1/FVC	rs7959413	12	96,147,836	T	C (0.47)	7.3×10^{-12}	5.2×10^{-6}	6.85	4.56	Agonist	0.005	*NTN4* intronic
PEF	rs10859942		96,161,207	T	C (0.49)	1.7×10^{-4}	2.0×10^{-5}	3.76	4.27	Agonist	0.011	*NTN4* intronic
FEV_1	rs59125351	15	96,144,157	T	G (0.24)	6.6×10^{-4}	6.9×10^{-11}	3.40	-6.52	Antagonist	0.035	Intergenic
PEF	rs8036214		96,145,329	T	C (0.49)	3.6×10^{-4}	5.0×10^{-5}	3.57	-4.06	Antagonist	0.021	Intergenic
PEF	rs16975396		96,158,705	T	G (0.24)	2.6×10^{-4}	3.9×10^{-10}	3.65	-6.26	Antagonist	0.015	Intergenic
FEV_1	rs8025061		96,176,965	A	G (0.43)	3.6×10^{-4}	1.2×10^{-5}	3.57	-4.38	Antagonist	0.023	Intergenic
PEF						2.4×10^{-5}	1.2×10^{-5}	4.22	-4.38	Antagonist	0.006	
PEF	rs16975446		96,177,806	G	A (0.33)	8.8×10^{-6}	2.3×10^{-5}	4.45	-4.23	Antagonist	0.011	Intergenic
FEV_1	rs3996842		96,178,859	C	T (0.33)	4.5×10^{-4}	2.9×10^{-5}	3.51	-4.18	Antagonist	0.023	Intergenic
FEV_1	rs4815366	20	25,049,909	T	G (0.34)	3.7×10^{-4}	5.6×10^{-5}	3.56	4.03	Agonist	0.026	Intergenic

Prioritization of gene candidates

To evaluate potential pleiotropic genes, we examined eQTL data from non-diseased human tissue (38), chromatin interactions identified in lung tissue (39), and protein physical and complex membership interactions (40, 41). Of the identified loci, the 5q31.3 LAM-pulmonary function shared variant rs7701443 maps in the first intron of the nuclear receptor subfamily 3 group C member 1 (*NR3C1*) and it is an eQTL for this gene (**Table S3**). NR3C1, also known as glucocorticoid receptor (GR), physically interacts with NR2F2, and this relationship influences the transcriptional regulation of defined gene targets (59). In addition, NR3C1 positively regulates the expression of another potential pleiotropic factor, *EPAS1*, in breast cancer cells under hypoxic conditions (60). Two other gene candidates, insulin-like growth factor-binding protein 3 (*IGFBP3*), defined by LAM-FEV_1 shared variant rs9648555, and disintegrin and metalloproteinase domain-containing protein 12 (*ADAM12*), defined by LAM-gastric cancer shared variant rs10901587, code for proteins that physically interact: ADAM12 cleaves IGFBP3 and this biochemical modification regulates IGF activity in regenerating and developing tissue, including cancer and pregnancy settings (61). The variants shared by LAM and pulmonary function detected in chromosome 8q23.3 might target the eukaryotic translation initiation factor 3 subunit H (*EIF3H*) gene, as proposed for the investigated associations with colorectal cancer risk (62).

Expression of candidate genes in LAM cells

A recent landmark study has depicted LAM single-cell gene expression profiles from four patients (8). This work defined a LAM^{core} expression signature that included *NR2F2* (8), as well as three of the candidates identified in our study: *ADAM12*, basonuclin 2 (*BNC2*) specificity protein 5 transcription factor (*SP5*) (**Table S3**). The identification of four LAM^{core} genes among 48 pleiotropic gene candidates (**Table S3**) was a higher proportion

than expected by chance (hypergeometric test of overlap P = 0.039). Examination of a second LAM single-cell study that analysed one tissue sample expanded the list of candidates to *IGFBP3* (7). Then, differential expression analysis between LAM and non-LAM cells using the first dataset (8) identified several gene candidates frequently detected and overexpressed in LAM cells. In addition to *NR2F2*, this analysis identified *EIF3H*, *IGFBP3*, and *NR3C1* as being linked to LAM cell profiles (**Table S4**). In turn, *EPAS1* was predicted to be underexpressed in LAM cells (**Table S4**).

Among the potential pleiotropic genes, we first evaluated *NR3C1* because its product has an established functional relationship with the primary LAM GWAS candidate, NR2F2 (59), and because of the key role played by glucocorticoids in breast cancer development and metastasis (63–65). Using RNA-seq data of 28 kidney AMLs (20, 43), a similar tumour entity to LAM, and of a large collection of 27 human cancer types (44), the expression level of *NR3C1* ranked relatively high in AMLs, just behind acute myeloid leukaemia and kidney renal clear cell carcinoma (**Figure 2a**). In addition, we found a significant (Fishers' exact test $P = 5.3 \times 10^{-10}$) gene overlap between an experimentally defined NR3C1-activity signature (66) and genes differentially expressed in LAM single cells (8) (**Figure 2b**). Then, immunohistochemistry assays in lung tissue from seven S-LAM patients confirmed substantial NR3C1 expression in all of them, with nuclear positivity in epithelioid and spindle-like diseased cells (**Figure 2c**). Co-staining with the LAM marker α-smooth muscle actin (αSMA) showed cellular colocalization with NR3C1 expression in lung nodules, although non-αSMA lung cells also presented nuclear expression of NR3C1 (**Figure 2d**).

Other candidate genes emerged from the exploration of eQTLs and genomic regulatory signals (**Table S3**), in combination with the suggestive evidence from other neoplasms. The netrin-4 (*NTN4*) gene —chromosome 12q22 LAM-pulmonary function shared associations (**Table 2**)— was identified as being the target of a breast cancer risk

association (67), and its product influences metastatic potential and angiogenesis (68, 69). Immunohistochemistry assays did not reveal NTN4 expression in LAM lung lesions, although the normal alveolar layer was positive (**Figure 3a**). Of the other candidates (**Table S3**), the contactin-2 (*CNTN2*) gene was prioritized based on criteria similar to those for *NTN4* (**Table S3**), and its product is linked to inflammatory conditions (70) and interacts with oestrogen receptor α (71). Since CNTN2 can be detected in body fluids (72), we measured its levels in the plasma of LAM patients and compared the results with those of healthy women and patients with related pulmonary diseases. Using ELISA assays, we identified significant overabundance of CNTN2 in LAM plasma relative to healthy women, and to Langerhans cell histiocytosis and Sjögren syndrome patients; however, we found no differences with respect to emphysema and systemic lupus erythematosus patients (**Figure 3b**). A comparison between LAM patients receiving and not receiving rapamycin treatment, and with low and high VEGF-D plasma levels (threshold of 800 pg/ml) revealed no significant differences (**Figure 3c**). Subsequent examination of single-cell transcriptomic data of the human lung (46) identified *CNTN2* expression exclusively in vascular smooth muscle cells (**Figure S3**). These cells also featured relatively high levels of expression of *NR2F2* and *NTN4*; in contrast, *NR3C1* was expressed in a wider range of lung cell types (**Figure S3**), consistent with the previous immunohistochemistry and immunohistofluorescence results (**Figure 2c,d**).

Evaluation of LAM causality

Following on the identified shared genetics and pleiotropic factors, we used Mendelian randomization methods (73) to evaluate causality. Analysis of LAM as outcome did not show significant associations with cancer, but suggested an association with FEV_1; this was supported by 245 genetic variants that depicted a LAM-FEV_1 negative correlation, as observed above, with no significant heterogeneity (**Table S5,** and **Figures S4** and

S5). However, a directionally test of causality showed inconsistent results with respect to the assumption that variance explained by the exposure should be greater that of the outcome (36) (**Table S5**). Then, evaluation of LAM as exposure suggested negative correlations with bladder and endometrial cancer risk, and a positive correlation with FEV_1; in these comparisons, the directionality assumption was fulfilled, but it was based on 8-9 variants (**Table S5**).

Discussion

This study identifies genetic variants that may concurrently influence LAM and cancer risk, or LAM and pulmonary function. Interestingly, two loci on chromosomes 4 and 9 are shared between kidney cancer risk and LAM risk, and the kidney is one potential tissue of origin for LAM (6, 8, 19, 48, 74–77). However, the gene candidates of these loci do not emerge as being linked to LAM single-cell transcriptome profiles (7, 8). In contrast, the genetic connection between LAM and pulmonary function proved to be more relevant in several analyses. There were indications of substantial shared genetic architecture in LAM and FVC, and PEF measures, and a trend in the same direction was observed for FEV_1. In all comparisons, the LAM-pulmonary function correlation was negative, which is consistent with the expectation that a given genetic variant that associates with greater pulmonary function would have a corresponding lower risk of LAM, and vice versa. It is of note that the data of pulmonary function came from a population-based study (24). In addition, the four pleiotropic gene candidates previously identified in a LAM single-cell signature (*ADAM12*, *BNC2*, *NR2F2* and *SP5*) (8) represented a higher proportion than would be expected by chance. However, the results of Mendelian randomization were not conclusive of a specific causal direction, likely because of the limited power and relative high variance of the LAM GWAS. These were also restrictions when examining genetic correlations, even if the HDL method (28) was used. There was an indication of

LAM being causal of reduced FEV_1, which was expected, but a causal effect in the opposite direction remains unclear.

The biological function of some of the identified pleiotropic candidates appears is consistent with LAM pathogenesis. NR3C1 (glucocorticoid receptor) can stimulate the transcriptional activity of NR2F2, which is the expected target of the primary LAM GWAS signal (20). The activity of both NR3C1 and NR2F2 interacts with hormone signalling in health and disease (78, 79) and this might be connected to the role of oestrogens in LAM pathogenesis (80). In addition, NR3C1 has been associated with breast cancer development and metastasis (63, 65), and may influence epithelial-to-mesenchymal transition, cell adhesion and tissue inflammation (53, 63, 81, 82). The confirmation of NR3C1 expression in LAM lung lesions and diseased cells is further evidence of a role for this factor in disease development, as well as in pulmonary function. Furthermore, the indication of increased NR3C1 activity in single-cell LAM transcriptome profiles might anticipate a therapeutic benefit from specific NR3C1 antagonists, as proposed for other hormonal cancers (83). It is of particular note that mifepristone, an antagonist of progesterone and glucocorticoid receptors yielded a preclinical benefit by inhibiting LAM tumorigenesis (84).

Other gene candidates were evaluated based on eQTL and genomic regulatory evidence. In a similar way to *NR3C1*, the identification of another gene linked to breast cancer risk, *NTN4* (67), is particularly intriguing. These observations might in turn be connected to our previous results of a potential higher incidence of breast cancer diagnoses among LAM patients (48, 49). We did not detect NTN4 expression in LAM-diseased cells, although a role for NTN4 influencing the tissue microenvironment cannot be excluded. Examining the other candidates, we found an overabundance of CNTN2 in LAM plasma, which, in addition to its potential for revealing information about disease risk, raises the possibility of establishing an independent biomarker of VEGF-D.

However, further analyses using samples collected at different times during the disease history are required to confirm these indications. Other pleiotropic gene candidates identified in this study might also be linked to LAM pathogenesis: IGF signalling and IGFBP2 function have been associated with LAM progression (85), though the expression of IGFBP3 in lung lesions is unclear (86); and LAM lung lesions show expression of matrix metallopeptidases (87), which also promote disease progression (88, 89). In addition, the LAM-FEV_1/FVC rs16937 and rs11240341 shared variants — which may target the transmembrane protein 81 (*TMEM81*) and/or retinoblastoma binding protein 5 (*RBBP5*) genes (90)— were previously associated with platelet measures in the general population (57) and, intriguingly, alteration of this blood component is associated with several inflammatory lung diseases (91).

Collectively, this study proposes a common aetiology between LAM and pulmonary function. This connection may be due to genes whose function is particularly relevant in the cell(s) of origin of LAM as well as lung tissue development and/or may indicate a cell origin of LAM that resides in the lung cell populations. The latter hypothesis appears to be consistent with the recent demonstration that *Tsc2* loss in the lung mesenchymal lineage causes LAM-like disease in mice (7). Thus, LAM might be an extreme phenotype of reduced lung function due to abnormal mTORC1-driven proliferation of mesenchymal-like cells. Particularly, NR3C1-glucocorticoid signalling regulates differentiation of proliferative mesenchymal progenitors into matrix fibroblasts (92), and these cells endorse synthesis of extracellular matrix and collagen during early lung development. The concept of altered mesenchymal cell differentiation and subsequent accumulation of extracellular matrix components would be consistent with the evolution of LAM lung pathology (93) and with identified pathway expression correlations with LAM^{core} (94). Intriguingly, pleural mesothelioma cell lines have transcriptome profiles relatively similar to those of LAM cells (94). In addition, the regulatory function of NR3C1 and NR2F2 may

be coordinated during lung development (92). Interestingly, GWAS signals of lung function were strongly predictive of chronic obstructive pulmonary disease (COPD) (24), which also shows extracellular matrix alterations and shares lymphatic vascular remodelling with LAM (95). Several of the lung function signals were also found to be associated with other respiratory traits, including asthma (24), which is frequently diagnosed in LAM women, and with an inflammatory multisystem disease predominantly affecting women, systemic lupus erythematosus, which can be presented with cystic lung disease (1, 2). Moreover, of the original 279 GWAS lung function signals (24), there were predicted 16 gene candidates that code for interactors of NR3C1, including the co-repressor NCOR1 (96) (**Table S6**).

The results of this study identify genetic factors and their molecular targets that may influence LAM development. However, as noted above, our study was limited by the relatively small sample size of the LAM GWAS, whose imputed summary statistics could also add noise to the results. Genetic studies of larger LAM cohorts are necessary to corroborate the findings and conclusively determine pleiotropy or causality with respect to pulmonary function and/or cancers. Likewise, studies of pleiotropic gene candidates may be warranted to better comprehend LAM aetiology and origin.

Acknowledgments

We wish to thank all the participants of the LAM GWAS for their contribution. We also thank all the investigators who were involved in the other GWASs analysed in this study for making their results available to us. This research was supported by “Asociación Española de LAM (AELAM)”, The LAM Foundation Seed Grant 2019, Carlos III Institute of Health grants PI18/01029, PI21/01306, and ICI19/00047 (co-funded by European Regional Development Fund (ERDF), “A way to build Europe”), Ministry of Economy and Competitivity grant SAF2017-88457-R, the “Generalitat de Catalunya” SGR 2017-449 and 2017-529, PERIS PFI-Salut SLT017-20-000076, and CERCA Program to IDIBELL and IGTP. X.F. is supported by the VEIS project (001-P-001647, ERDF Operational Programme of Catalonia 2014-2020, co-funded by European Regional Development Fund (ERDF), “A way to build Europe”).

Figure legends

Figure 1

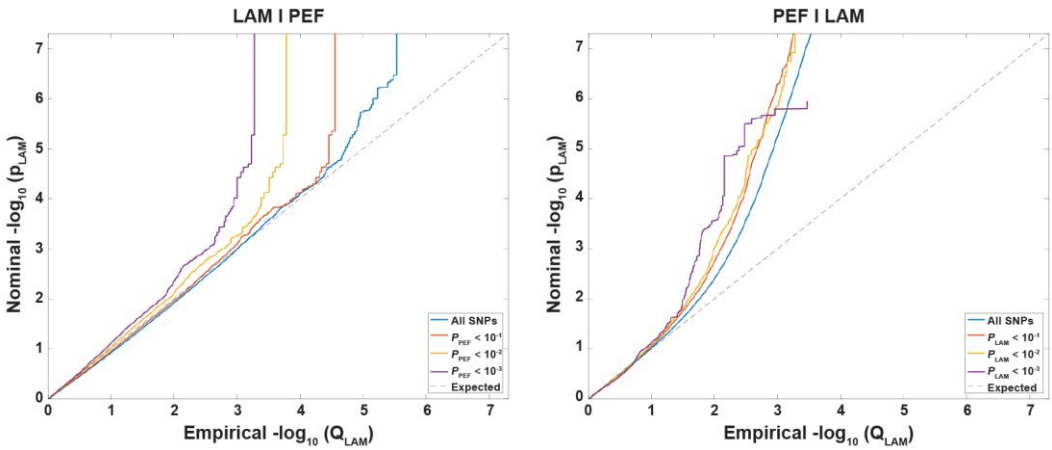

Figure 1. Stratified Q-Q plots including LAM and PEF GWAS results. Left panel, Q-Q plot (nominal versus empirical -$\log_{10}$ *P* values, corrected for inflation) conditioning LAM on PEF; and right panel, Q-Q plot conditioning PEF on LAM. Leftwards deflection from the null distribution of the observed *P* value as the thresholds become more stringent, indicates genetic overlap between the two traits.

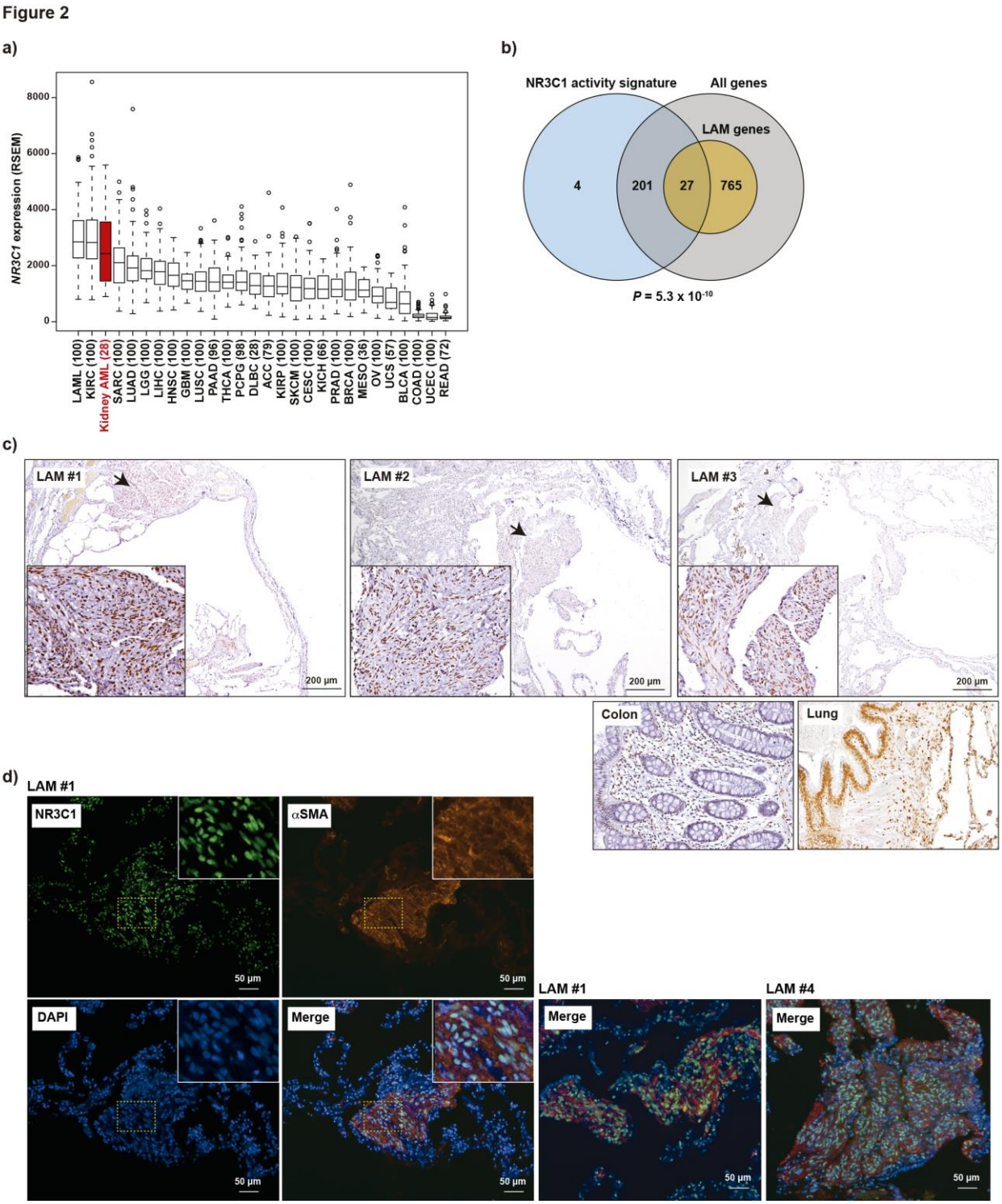

Figure 2. *NR3C1* gene expression in AML tumours and NR3C1 protein expression in LAM lung lesions. a) Comparison of *NR3C1* expression in kidney AMLs (red font) with other neoplasms (TCGA data: 2,463 tumors of 27 histological types). Cancer abbreviations in descending order: LAML: acute myeloid leukaemia; KIRC: kidney renal clear cell carcinoma; SARC: sarcoma; LUAD: lung adenocarcinoma; LGG: low-grade

glioma; LIHC: liver hepatocellular carcinoma; HNSC: head and neck squamous cell carcinoma; GBM: glioblastoma multiforme; LUSC: lung squamous cell carcinoma; PAAD: pancreatic adenocarcinoma; THCA: thyroid carcinoma; PCPG: phaeochromocytoma and paraganglioma; DLBC: lymphoid neoplasm diffuse large B-cell lymphoma; ACC: adrenocortical carcinoma; KIRP: kidney renal papillary cell carcinoma; SKCM: skin cutaneous melanoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; KICH: kidney chromophobe; PRAD: prostate adenocarcinoma; BRCA: breast invasive carcinoma; MESO: mesothelioma; OV: ovarian serous cystadenocarcinoma; UCS: uterine carcinosarcoma; BLCA: bladder urothelial carcinoma; COAD: colon adenocarcinoma; UCEC: uterine corpus endometrial carcinoma; and READ: rectum adenocarcinoma. The numbers in parentheses are the sample sizes of the indicated cancer types. The average, interquartile range and 95% range are shown for each setting, with outliers indicated by circles. Gene expression is shown as RSEM (RNA sequencing by expectation maximization) values. b) Venn diagram showing the identity overlap (n = 27) between genes identified in the NR3C1-activity signature and genes differentially expressed in LAM single cells. c) Representative images from immunohistochemistry assays for the detection of NR3C1 expression in LAM lung lesions of three patients (LAM #1-3). The arrows indicate the magnified lesion areas in the insets: in magnified lung nodules, epithelioid and spindle-like diseased cells are apparent from the observed nuclear shapes of positive NR3C1 staining. The positive control results from colon tissue are also shown, as well as those from normal lung tissue showing positivity in the alveolar epithelium, and luminal and basal layers of the bronchioles. d) Representative images of immunohistofluorescence detection and colocalization of NR3C1 and αSMA in LAM lung lesions; nuclei stained blue with DAPI (merged). Lung nodules of three LAM patients are shown.

Figure 3

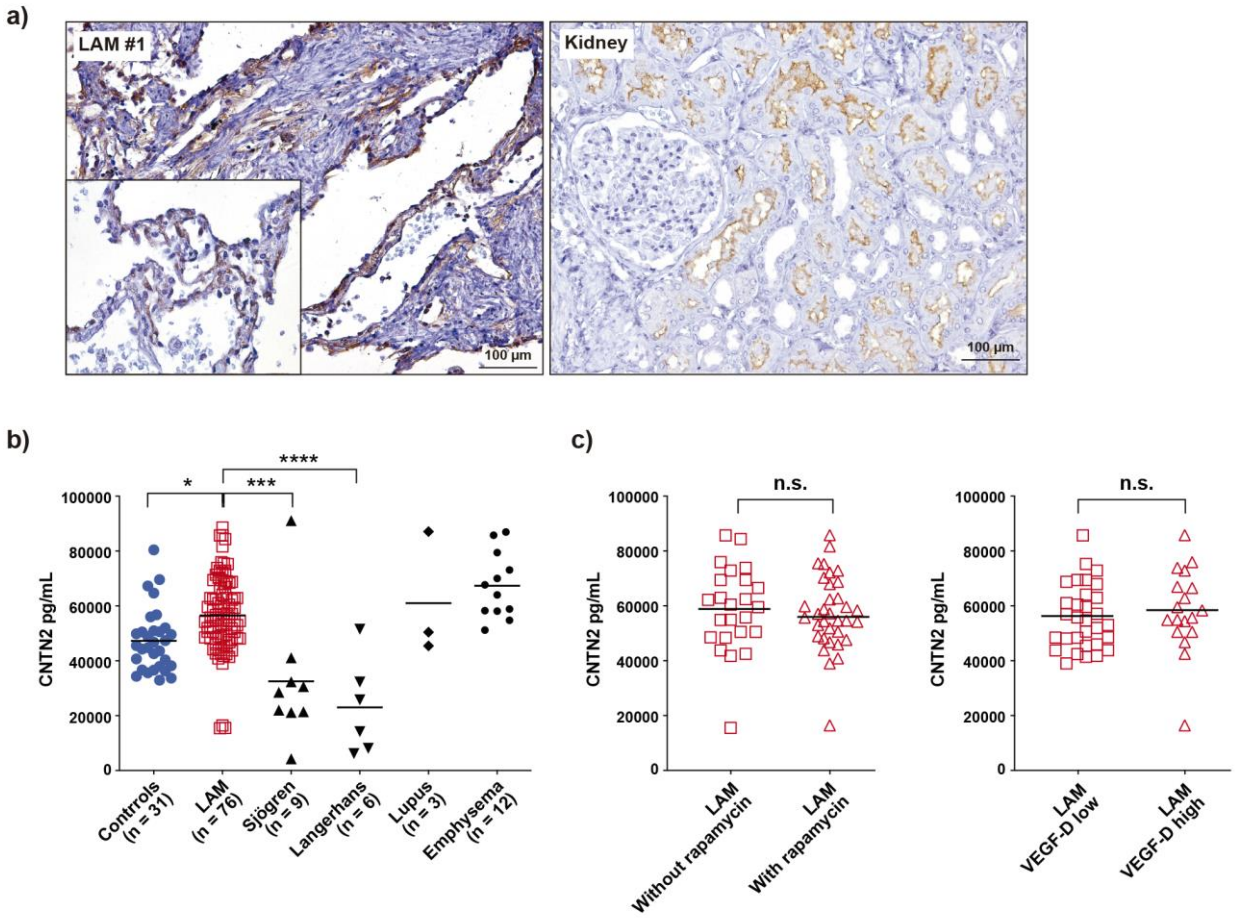

Figure 3. Evaluation of additional pleiotropic factors. a) Representative images from immunohistochemistry assays for detecting NTN4 expression in LAM lung lesions. Lung nodules appear negative, while the alveolar layer is positive. The positive control of kidney tissue is shown. b) Overabundance of CNTN2 in LAM plasma relative to healthy women and two related pulmonary diseases, as indicated. The number of samples analysed in each setting (n) is indicated. Asterisks indicate significant differences based on two-sided Mann-Whitney tests (*P < 0.05, ***P < 0.001 and ****P < 0.0001). Average values are indicated with horizontal black lines. c) Absence of significant differences (not significant; n.s.) of CNTN2 plasma levels between LAM patients receiving and not receiving rapamycin treatment (left panel), and between LAM patients with high and low VEGF-D plasma levels (right panel).

Supplementary figure legends

Figure S1

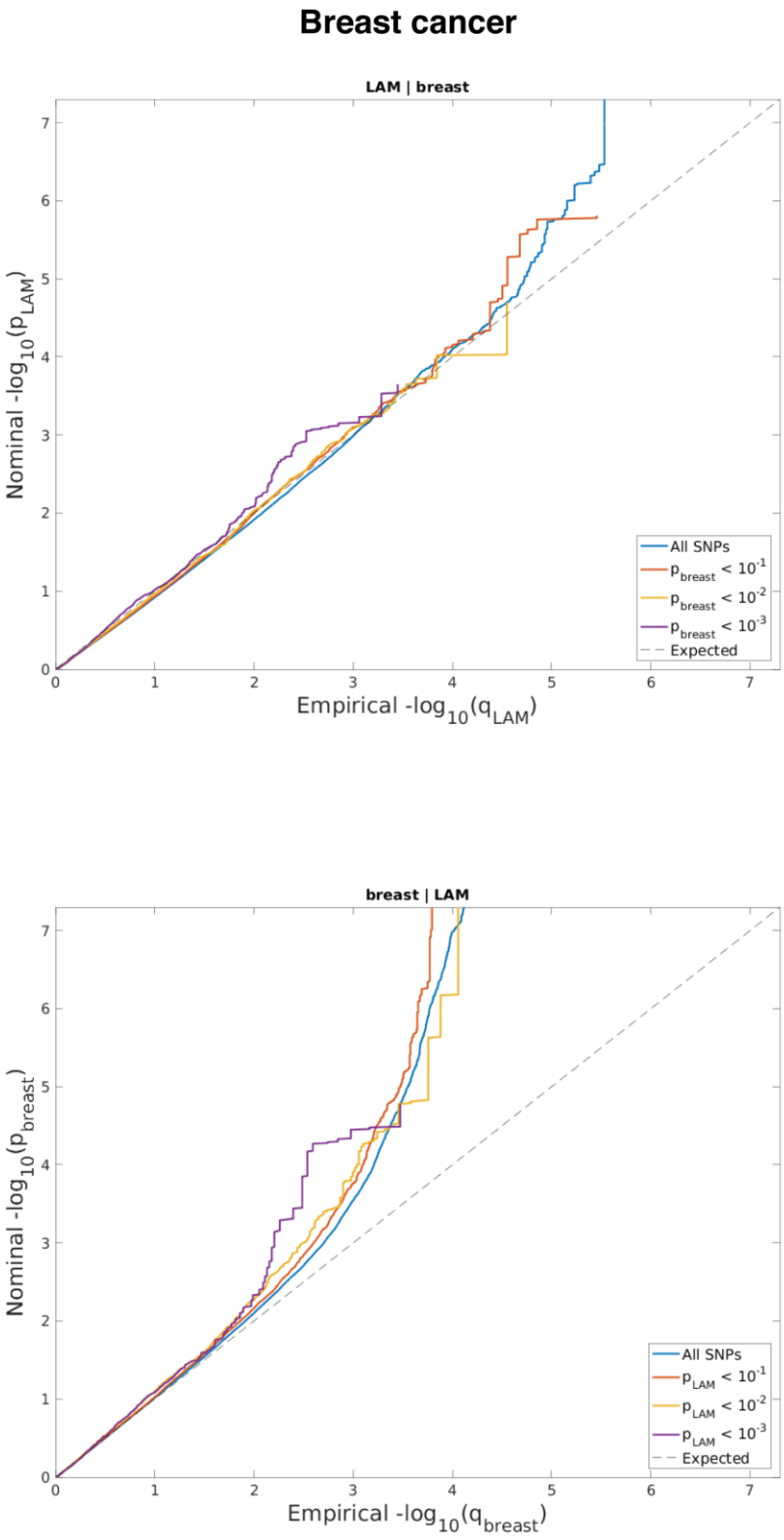

Figure S1. Stratified Q-Q plots including LAM and cancer GWAS results. The Q-Q plots (nominal versus empirical $-\log_{10}$ *P* values, corrected for inflation) conditioning for LAM or each cancer type are shown consecutively.

Figure S2

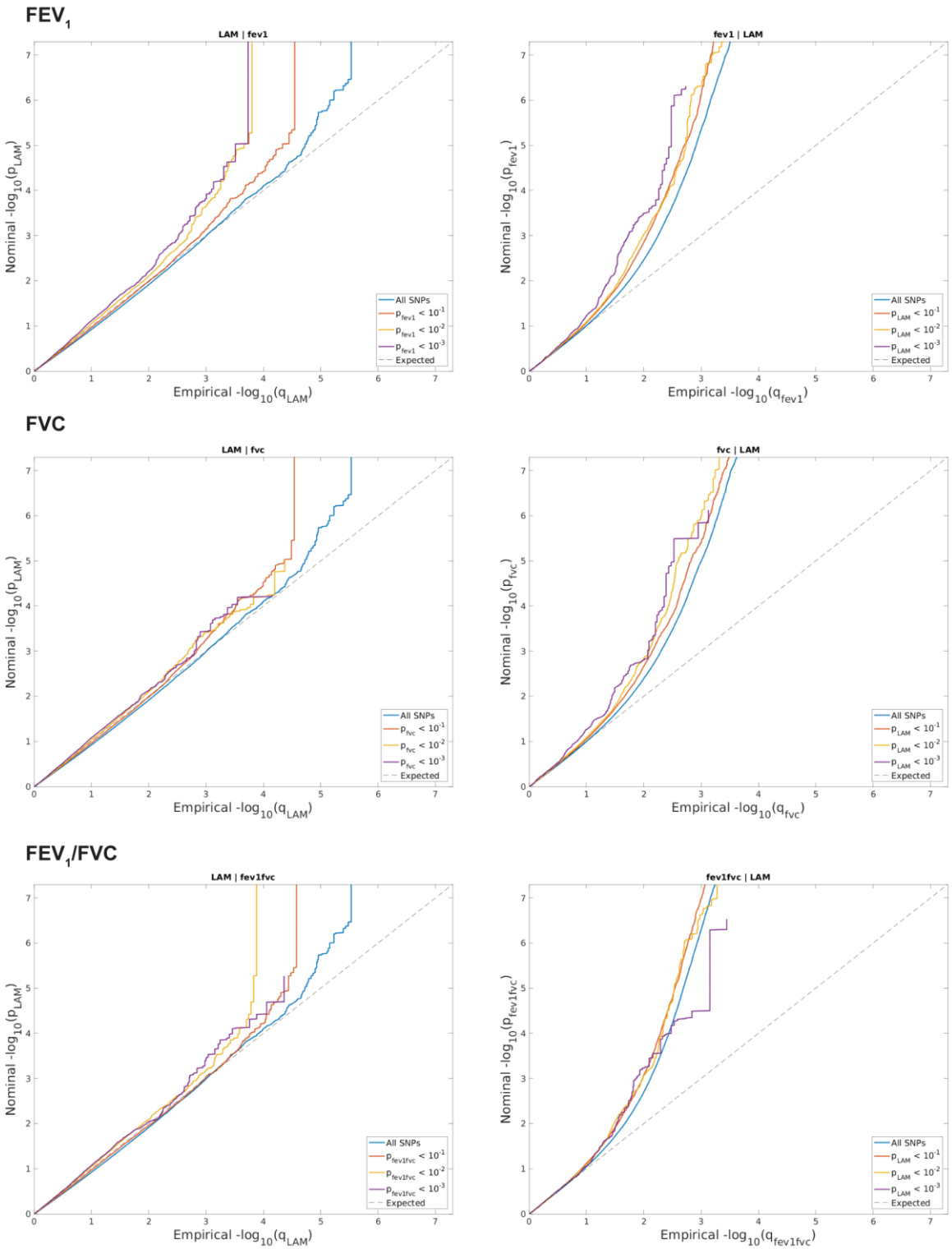

Figure S2. Stratified Q-Q plots including LAM and FEV_1, FVC and FEV_1/FVC ratio GWAS results. The Q-Q plots (nominal versus empirical -$\log_{10}$ *P* values, corrected for inflation) conditioning for LAM or each spirometry measure are shown consecutively.

Figure S3

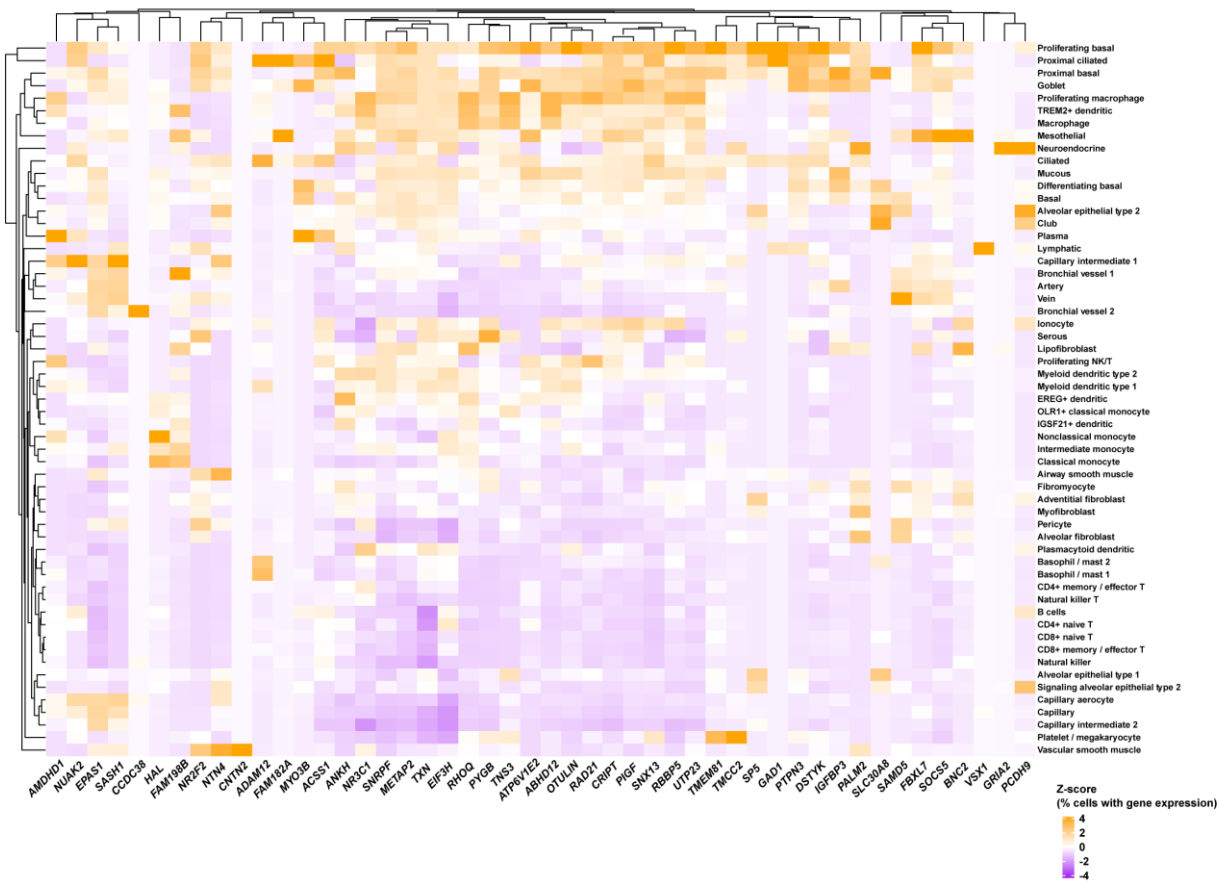

Figure S3. Expression levels of candidate pleiotropic genes across human lung cell types. Unsupervised hierarchical clustering of gene expression across cell types. The values are equivalent to the percentage of cells showing expression of a given gene.

Figure S4

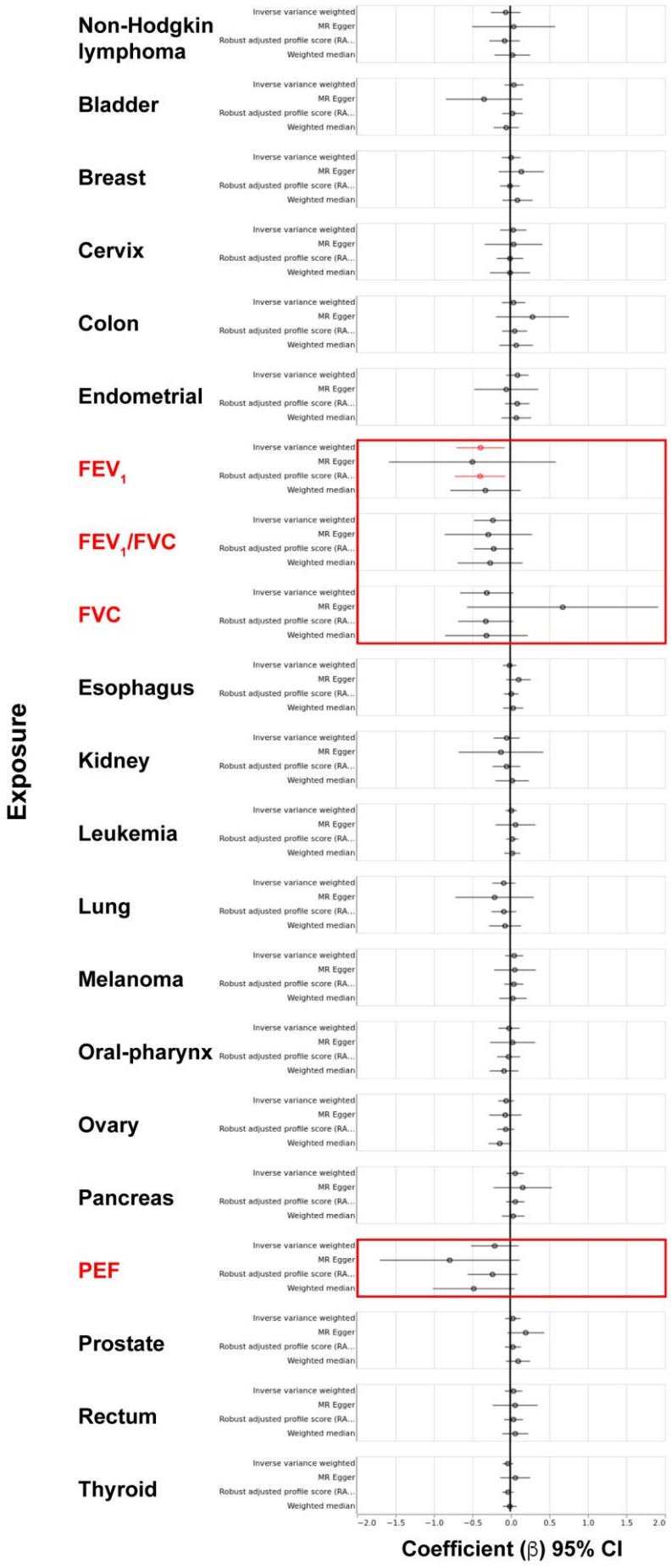
Exposure
Non-Hodgkin lymphoma
Bladder
Breast
Cervix
Colon
Endometrial
FEV1
FEV1/FVC
FVC
Esophagus
Kidney
Leukemia
Lung
Melanoma
Oral-pharynx
Ovary
Pancreas
PEF
Prostate
Rectum
Thyroid
Inverse variance weighted
MR Egger
Robust adjusted profile score (RA...
Weighted median
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Coefficient (β) 95% CI

Figure S4. Genetic effects of defined exposures on outcome, LAM. Forest plot showing the inverse variance-weighted, weighted median, MR-Egger and MR-RAPS estimates, and 95% CIs, of each exposure analysis. Red letters and rectangles indicate the results of pulmonary function exposures.

Figure S5

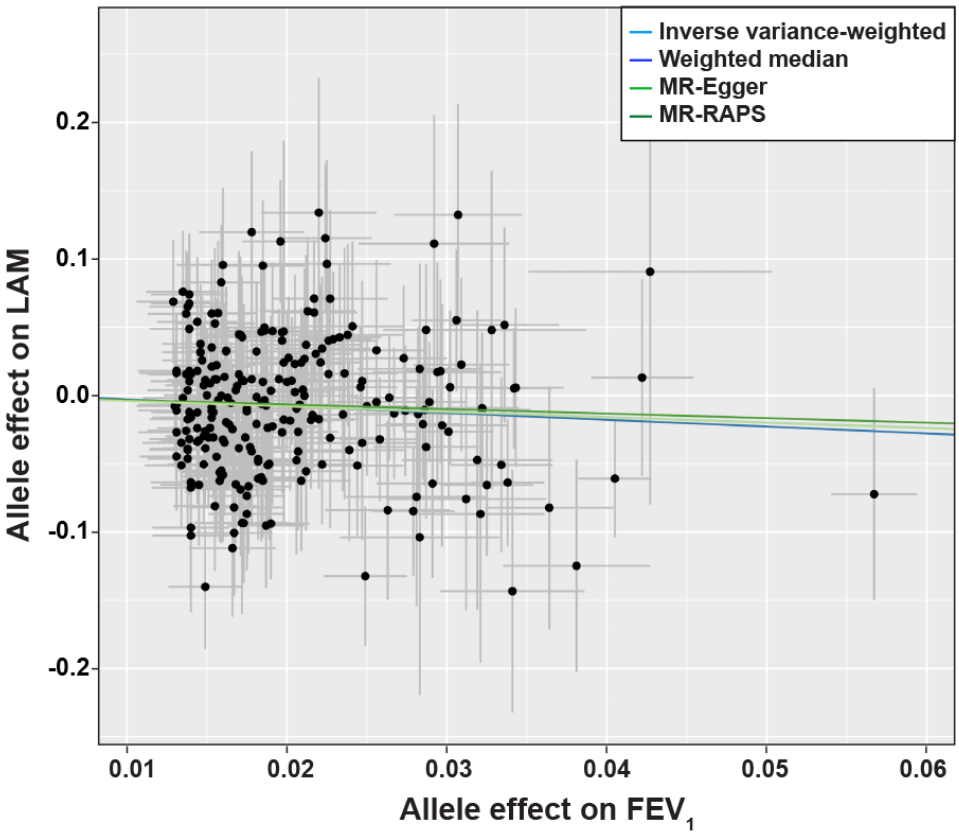

Figure S5. Potential causal connection between LAM and pulmonary function. Scatter plot showing the associations of the genetic effects on outcome (LAM, log odds ratio) against the effects on the exposure (FEV_1, log hazard ratio). The inverse variance-weighted, weighted median, MR-Egger, and MR-RAPS estimates are represented by lines as indicated in the inset.

Supplementary tables

Table S1. GWAS data sources.

Disease/trait	Domain	Source	Cases (*n*)	Controls (*n*)	Total (*n*)
LAM	Respiratory	https://doi.org/10.1183/13993003.00329-2019	426	852	1278
PEF	Respiratory	https://doi.org/10.1038/s41588-018-0321-7	NA	NA	400102
FEV1	Respiratory	https://doi.org/10.1038/s41588-018-0321-7	NA	NA	400102
FVC	Respiratory	https://doi.org/10.1038/s41588-018-0321-7	NA	NA	400102
FEV1/FVC	Respiratory	https://doi.org/10.1038/s41588-018-0321-7	NA	NA	400102
Bladder cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	2242	410350	412592
Breast cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	17881	410350	428231
Cervix cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	6563	410350	416913
Colon cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	3793	410350	414143
Endometrium cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	2037	410350	412387
Esophagus/Stomach cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	1091	410350	411441
Kidney cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	1338	410350	411688
Lymphocytic Leukemia	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	852	410350	411202
Lung cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	2485	410350	412835
Melanoma	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	6777	410350	417127
Non-Hodgkin's Lymphoma	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	2400	410350	412750
Oral Cavity/Pharynx cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	1223	410350	411573
Ovary cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	1259	410350	411609
Pancreas cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	663	410350	411013
Prostate cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	10792	410350	421142
Rectum cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	2091	410350	412441
Thyroid cancer	Neoplasms	https://doi.org/10.1038/s41467-020-18246-6	762	410350	411112

Table S1. Metadata of the GWAS summary statistics.

Table S2. Heritability and genetic correlation between LAM and cancer risk, and LAM and spirometry measures.

Trait (cancer risk or pulmonary function measure)	Trait heritability	Heritability (se)	Genetic covariance	Genetic covariance (se)	Genetic correlation	Genetic correlation (se)	Genetic correlation (95% CI)	*P* value
FVC	0.1604	0.0051	-0.0295	0.0119	-0.0737	0.0297	(-0.044,-0.1034)	1.31E-02
PEF	0.1461	0.0071	-0.0384	0.0176	-0.1004	0.0460	(-0.0544,-0.1464)	2.90E-02
Cervix	0.0080	0.0017	-0.0288	0.0170	-0.3224	0.1860	(-0.1364,-0.5084)	8.30E-02
FEV1	0.1636	0.0054	-0.0245	0.0148	-0.0607	0.0365	(-0.0242,-0.0972)	9.67E-02
Bladder	0.0014	0.0009	-0.0223	0.0106	-0.5977	0.3904	(-0.2073,-0.9881)	1.26E-01
Leukemia	0.0001	0.0005	0.0086	0.0128	0.7322	0.5001	(1.2323,0.2321)	1.43E-01
Rectum	0.0008	0.0008	-0.0064	0.0106	-0.2273	0.3893	(0.162,-0.6166)	5.59E-01
Colon	0.0041	0.0011	-0.0043	0.0089	-0.0674	0.1412	(0.0738,-0.2086)	6.33E-01
Breast	0.0240	0.0049	-0.0077	0.0169	-0.0498	0.1092	(0.0594,-0.159)	6.48E-01
Lung	0.0027	0.0009	-0.0051	0.0115	-0.0982	0.2242	(0.126,-0.3224)	6.61E-01
Non-Hodgkin lymphoma	0.0012	0.0010	-0.0053	0.0119	-0.1509	0.3605	(0.2096,-0.5114)	6.75E-01
Melanoma	0.0069	0.0018	0.0045	0.0133	0.0542	0.1621	(0.2163,-0.1079)	7.38E-01
Prostate	0.0271	0.0049	-0.0036	0.0171	-0.0216	0.1037	(0.0821,-0.1253)	8.35E-01
FEV1/FVC	0.1680	0.0080	-0.0004	0.0176	-0.0009	0.0430	(0.0421,-0.0439)	9.83E-01
Ovary	0.0000	0.0000	-0.0164	0.0159	Inf	NA	NA	NA
Pancreas	0.0000	0.0000	0.0007	0.0125	Inf	NA	NA	NA
Gastric and esophageal	0.0000	0.0000	0.0200	0.0128	Inf	NA	NA	NA
Thyroid	0.0000	0.0000	0.0151	0.0113	Inf	NA	NA	NA
Endometrial	0.0000	0.0000	-0.0036	0.0156	Inf	NA	NA	NA
Kidney	0.0000	0.0000	-0.0124	0.0115	Inf	NA	NA	NA
Oropharyngeal	0.0000	0.0000	-0.0239	0.0129	Inf	NA	NA	NA

se: standard error; CI, confidence interval

Table S2. Genetic correlation between LAM and cancer risk, and LAM and spirometry measures.

Table S3. Shared variants between LAM and cancer or pulmonary function, and candidate pleiotropic genes.

SNP ID	Chromosome	Position (bp, hg19)	Reference	Alternative	Locus gene(s)	GTEx pan-tissue	GTEx eQTL nominal (blood vessel, lung, kidney, and u	GWAS Catalog	4C lung interactions and eQTL	LAMcore (G	LAM vs mesenchymal (Guo et al.)	LAM (Obraztsova et al.)
rs11240341	1	205015284	T	C	*CNTN2*	*CNTN2, DSTYK, NUAK2, RBBP*	NA	Mean platelet volume, https://www.ebi.ac.uk/gwas/publications/27863252; Schizophrenia, https://www.ebi.ac.uk/gwas/studies/GCST004946	*CNTN2, DSTYK, NUAK2, TMCC2*	No	No	No
rs16937	1	205035455	A	G	*CNTN2*	*CNTN2, DSTYK, NUAK2, RBBP*	NA	Mean platelet volume, https://www.ebi.ac.uk/gwas/publications/27863252; Schizophrenia, https://www.ebi.ac.uk/gwas/studies/GCST004946	*CNTN2, DSTYK, NUAK2, TMCC2*	No	No	No
rs13410076	2	46815961	C	T	*PIGF*	*ATP6V1E2, CRIPT, PIGF*	*ATP6V1E2* (aorta, coronary, lymphocytes, tibial, uterus), (	NA	*PRKCE, EPAS1, TMEM247, ATP*	No	*CRIPT*	No
rs4668267	2	171360896	C	T	*MYO3B*	*GAD1, LINC01124, SP5*	*GAD1* (coronary, tibial)	Lobe attachment; http://www.ebi.ac.uk/efo/EFO_0007667	*GAD1, MYO3B, UBR3, GORASP*	*SP5*	*SP5*	No
rs17036640	4	158605184	G	A	*GRIA2* (> 300 kb), *FAM198B* (> 300 kb)	*RP11-364P22.2*	*FAM198B* (aorta)	NA	*GRIA2, GASK1B/FAM198B, TM*	No	No	No
rs4512050	4	158597491	G	A	*GRIA2* (> 300 kb), *FAM198B* (> 300 kb)	*RP11-364P22.2*	*FAM198B* (aorta)	NA		No	No	No
rs7701443	5	142792650	G	A	*NR3C1*	*NR3C1*	*NR3C1* (aorta)	NA	*NR3C1, ARHGAP26, HMHB1*	No	No	No
rs835069	5	14984116	T	C	*ANKH* (> 200 kb)	NA	NA	NA	*ANKH, FAM105B/OTULIN, TRI*	No	No	No
rs3861451	6	148320047	C	T	*SAMD5* (> 500 kb), *SASH1* (> 200 kb)	*SAMD5*	*SAMD5* (coronary, lung, tibial)	NA	*SASH1*	No	No	No
rs9648555	7	46656869	G	A	*LOC730338, IGFBP3* (> 500 kb), *TNS3* (> 500 kb)	NA	NA	NA		No	No	*IGFBP3*
rs2537572	7	17771780	G	A	*SNX13*	*RP11-511H23.2*	*RP11-511H23.2* (kidney)	NA	*SNX13, HDAC9, AHR*	No	No	No
rs12542425	8	117642345	T	C	*EIF3H, UTP23*	*UTP23*	NA	NA	*SLC30A8*	No	No	No
rs34672734	8	117805281	G	A	*UTP23, RAD21*	*SLC30A8, UTP23*	*RAD21* (aorta)	NA	*EIF3H, SLC30A8, LINC00536*	No	No	No
rs17663673	8	117650006	C	T	*EIF3H*	*UTP23*	*EIF3H* (aorta, lymphocytes), *UTP23* (aorta, lung, tibial)	NA	*EIF3H, SLC30A8, LINC00536*	No	No	No
rs7839361	8	117645687	C	T	*EIF3H*	NA	NA	NA		No	No	No
rs2511654	8	117623003	C	T	*EIF3H*	NA	NA	NA		No	No	No
rs2146084	9	112505617	A	G	*PALM2*	*PALM2*	*PALM2* (tibial)	NA	*PALM2, TXN, PTPN3*	No	No	No
rs4961722	9	16529174	C	T	*BNC2*	NA	*BNC2* (uterus)	NA	*BNC2, C9orf92*	*BNC2*	No	No
rs10901587	10	128001759	T	C	*ADAM12*	*ADAM12*	*ADAM12* (uterus)	NA	*ADAM12, C10orf90*	*ADAM12*	No	No
rs10859942	12	96161207	C	T	*NTN4*	*AMDHD1, HAL, NTN4, RP11-5*	*AMDHD1* (lung), *NTN4* (artery tibial, uterus), *SNRPF* (ar	NA	*METAP2,*	No	No	No
rs7959413	12	96147836	C	T	*NTN4*	*HAL, SNRPF, RP11-536G4.2*	*SNRPF* (artery tibial)		*NTN4, SNRPE*	No	No (*SNRP* genes included)	No (*SNRP* genes included)
rs9528802	13	65235556	C	T	*PCDH9* (> 1 Mb)	NA	NA	NA		No	No	No
rs8025061	15	96176965	G	A	*NR2F2* (> 500 kb)	NA	*NR2F2* (aorta)	NA		*NR2F2*	No	No
rs16975446	15	96177806	A	G	*NR2F2* (> 500 kb)	NA	*NR2F2* (aorta)	NA		*NR2F2*	No	No
rs16975396	15	96158705	G	T	*NR2F2* (> 500 kb)	NA	NA	NA		*NR2F2*	No	No
rs8036214	15	96145329	C	T	*NR2F2* (> 500 kb)	NA	NA	NA		*NR2F2*	No	No
rs3996842	15	96178859	T	C	*NR2F2* (>500kb)	NA	*NR2F2* (aorta)	NA		*NR2F2*	No	No
rs59125351	15	96144157	G	T	*NR2F2* (> 500 kb)	NA	NA	NA		*NR2F2*	No	No
rs4815366	20	25049909	G	T	*ACSS1, VSX1*	*ABHD12, ACSS1, FAM182A, PI*	*ABHD12* (aorta, coronary, lung, tibial), *ACSS1* (coronary	NA	*ABHD12, GINS1, NINL, ENTPD6*	No	No	No

Table S4. Expression analysis of candidate pleiotropic genes using dataset of Guo et al.

Gene	LAM cells detected	LAM cells non-detected	LAM cells %	non-LAM cells detected	non-LAM cells non-detected	non-LAM cells %	LAM cells mean expression	non-LAM cells mean expression	Wilcoxon test P value	Wilcoxon FDR	Z-test P value	Z-test FDR
METAP2	41	92	30.83	9220	11705	44.06	1.58	0.92	1.44E-08	4.02E-07	2.90E-03	6.64E-03
EIF3H	70	63	52.63	15459	5466	73.88	1.80	1.37	8.68E-07	1.22E-05	4.98E-08	5.98E-07
SNRPF	48	85	36.09	12330	8595	58.92	1.56	1.02	1.73E-06	1.62E-05	1.57E-07	1.25E-06
NR2F2	35	98	26.32	860	20065	4.11	1.38	0.97	1.13E-05	7.88E-05	1.61E-35	3.86E-34
CRIPT	24	109	18.05	4512	16413	21.56	1.20	0.72	1.54E-05	8.60E-05	3.80E-01	6.08E-01
TNS3	5	128	3.76	3782	17143	18.07	1.65	0.68	2.18E-03	1.02E-02	3.02E-05	1.32E-04
ABHD12	19	114	14.29	5769	15156	27.57	1.08	0.73	4.76E-03	1.90E-02	8.90E-04	2.67E-03
RHOQ	15	118	11.28	4873	16052	23.29	1.15	0.73	6.62E-03	2.32E-02	1.54E-03	3.89E-03
RAD21	15	118	11.28	6080	14845	29.06	1.21	0.84	3.72E-02	1.16E-01	1.03E-05	4.94E-05
NR3C1	20	113	15.04	7440	13485	35.56	1.12	0.87	5.36E-02	1.50E-01	1.29E-06	7.76E-06
ANKH	11	122	8.27	1425	19500	6.81	1.22	0.79	6.61E-02	1.68E-01	6.22E-01	9.32E-01
IGFBP3	7	126	5.26	286	20639	1.37	1.49	1.09	7.49E-02	1.75E-01	5.55E-04	1.78E-03
FBXL7	5	128	3.76	174	20751	0.83	0.68	1.20	1.10E-01	2.20E-01	1.41E-03	3.76E-03
DSTYK	7	126	5.26	1212	19713	5.79	0.85	0.72	1.07E-01	2.20E-01	9.41E-01	1.00E+00
SP5	15	118	11.28	46	20879	0.22	1.02	0.80	1.31E-01	2.29E-01	1.63E-115	7.82E-114
PIGF	8	125	6.02	3448	17477	16.48	0.98	0.66	1.29E-01	2.29E-01	1.75E-03	4.19E-03
NTN4	5	128	3.76	2930	17995	14.00	1.36	1.00	2.13E-01	3.50E-01	1.06E-03	2.99E-03
TXN	86	47	64.66	17371	3554	83.02	2.17	2.08	2.44E-01	3.59E-01	4.05E-08	5.98E-07
FAM198B	5	128	3.76	656	20269	3.14	1.12	0.82	2.40E-01	3.59E-01	8.71E-01	1.00E+00
ADAM12	9	124	6.77	297	20628	1.42	0.76	0.79	4.98E-01	6.65E-01	1.81E-06	9.64E-06
PTPN3	3	130	2.26	1076	19849	5.14	0.59	0.73	4.91E-01	6.65E-01	1.91E-01	3.27E-01
EPAS1	17	116	12.78	7363	13562	35.19	0.91	1.02	5.38E-01	6.81E-01	1.11E-07	1.07E-06
RBBP5	4	129	3.01	1324	19601	6.33	0.80	0.69	5.59E-01	6.81E-01	1.64E-01	3.03E-01
SNX13	4	129	3.01	2881	18044	13.77	0.73	0.68	6.52E-01	7.60E-01	5.18E-04	1.78E-03
BNC2	11	122	8.27	378	20547	1.81	0.97	1.04	8.14E-01	8.14E-01	2.04E-07	1.40E-06
UTP23	3	130	2.26	2780	18145	13.29	1.04	0.67	7.57E-01	8.14E-01	3.00E-04	1.11E-03
SOCS5	9	124	6.77	1078	19847	5.15	0.71	0.72	7.86E-01	8.14E-01	5.20E-01	8.06E-01
ACSS1	10	123	7.52	1663	19262	7.95	0.64	0.74	7.28E-01	8.14E-01	9.83E-01	1.00E+00
PYGB	2	131	1.50	2765	18160	13.21	1.64	0.66	NA	NA	1.15E-04	4.61E-04
SASH1	2	131	1.50	1835	19090	8.77	0.86	0.80	NA	NA	5.02E-03	1.09E-02
PCDH9	0	133	0.00	1066	19859	5.09	NA	0.91	NA	NA	1.34E-02	2.80E-02
FAM105B	2	131	1.50	1462	19463	6.99	0.99	0.70	NA	NA	2.10E-02	4.21E-02
PALM2	2	131	1.50	57	20868	0.27	0.73	0.71	NA	NA	6.36E-02	1.22E-01
ATP6V1E2	1	132	0.75	670	20255	3.20	0.49	0.68	NA	NA	1.75E-01	3.11E-01
SAMD5	0	133	0.00	398	20527	1.90	NA	0.85	NA	NA	1.98E-01	3.28E-01
HAL	0	133	0.00	147	20778	0.70	NA	0.67	NA	NA	6.54E-01	9.52E-01
NUAK2	1	132	0.75	332	20593	1.59	0.66	0.76	NA	NA	6.74E-01	9.52E-01
AMDHD1	0	133	0.00	74	20851	0.35	NA	0.74	NA	NA	1.00E+00	1.00E+00
CCDC38	0	133	0.00	4	20921	0.02	NA	1.04	NA	NA	1.00E+00	1.00E+00
CNTN2	0	133	0.00	3	20922	0.01	NA	0.46	NA	NA	1.00E+00	1.00E+00
FAM182A	0	133	0.00	18	20907	0.09	NA	0.83	NA	NA	1.00E+00	1.00E+00
GAD1	0	133	0.00	38	20887	0.18	NA	0.55	NA	NA	1.00E+00	1.00E+00
GRIA2	0	133	0.00	5	20920	0.02	NA	0.97	NA	NA	1.00E+00	1.00E+00
MYO3B	0	133	0.00	28	20897	0.13	NA	0.89	NA	NA	1.00E+00	1.00E+00
SLC30A8	0	133	0.00	3	20922	0.01	NA	0.56	NA	NA	1.00E+00	1.00E+00
TMCC2	0	133	0.00	63	20862	0.30	NA	0.77	NA	NA	1.00E+00	1.00E+00
TMEM81	0	133	0.00	47	20878	0.22	NA	0.88	NA	NA	1.00E+00	1.00E+00
VSX1	0	133	0.00	2	20923	0.01	NA	0.54	NA	NA	1.00E+00	1.00E+00

Table S5. Results of MR analysis.

			MR-EGGER							Inverse Variance Weighted							MR-RAPS				MR Steiger Directionality Test			
Exposure	Outcome	SNPs (n)	Coefficient	Standard error	Coefficient 95% CI	P value	Heterogeneity (Q)	Heterogeneity (df)	Heterogeneity (P value)	Coefficient	Standard error	Coefficient 95% ci	P value	Heterogeneity (Q)	Heterogeneity (df)	Heterogeneity (P value)	Coefficient	Standard error	Coefficient 95% ci	P value	Variance explained exposure	Variance explained outcome	Correct causal direction	P value
Bladder	LAM	24	-0.349	0.254	(-0.847,-0.094)	1.84E-01	12.24	22	9.52E-01	0.040	0.063	(0.163,-0.024)	5.32E-01	14.73	23	9.04E-01	0.020	0.068	(0.154,-0.048)	7.71E-01	0.002	0.018	False	1.00E-02
Breast	LAM	84	0.134	0.149	(-0.159,0.283)	3.72E-01	69.89	82	8.27E-01	0.003	0.062	(0.125,-0.06)	9.67E-01	70.83	83	8.27E-01	-0.013	0.066	(0.116,-0.079)	8.44E-01	0.019	0.088	False	2.94E-07
Cervix	LAM	20	0.035	0.191	(-0.339,0.226)	8.57E-01	21.64	18	2.48E-01	0.031	0.087	(0.202,-0.056)	7.24E-01	21.65	19	3.02E-01	-0.012	0.087	(0.159,-0.099)	8.93E-01	0.005	0.032	False	8.40E-04
Colon	LAM	25	0.280	0.241	(-0.193,0.522)	2.57E-01	17.50	23	7.84E-01	0.033	0.078	(0.185,-0.044)	6.68E-01	18.67	24.00	7.69E-01	0.048	0.083	(0.211,-0.035)	5.61E-01	0.002	0.030	False	1.24E-04
Endometrial	LAM	20	-0.061	0.212	(-0.476,0.151)	7.78E-01	12.12	18	8.41E-01	0.084	0.074	(0.229,0.01)	2.54E-01	12.65	19	8.56E-01	0.081	0.080	(0.238,0.001)	3.12E-01	0.003	0.021	False	7.39E-03
Esophagus	LAM	17	0.099	0.082	(-0.06,0.181)	2.42E-01	12.40	15	6.48E-01	-0.016	0.045	(0.072,-0.061)	7.21E-01	15.29	16	5.03E-01	0.003	0.049	(0.099,-0.046)	9.48E-01	0.003	0.016	False	3.61E-02
FEV1	LAM	245	-0.500	0.553	(-1.584,0.053)	3.67E-01	223.07	243	8.16E-01	-0.395	0.160	(-0.081,-0.555)	1.38E-02	223.11	244	8.27E-01	-0.402	0.167	(-0.074,-0.569)	1.63E-02	0.037	0.226	False	4.42E-25
FEV1/FVC	LAM	272	-0.293	0.290	(-0.86,-0.003)	3.13E-01	246.31	270	8.47E-01	-0.234	0.125	(0.012,-0.359)	6.21E-02	246.36	271	8.56E-01	-0.223	0.130	(0.032,-0.353)	8.71E-02	0.059	0.362	False	3.67E-24
FVC	LAM	219	0.672	0.634	(-0.57,1.306)	2.90E-01	230.06	217	2.59E-01	-0.314	0.177	(0.033,-0.491)	7.59E-02	232.84	218	2.34E-01	-0.327	0.183	(0.032,-0.51)	7.39E-02	0.030	0.241	False	1.63E-31
Kidney	LAM	8	-0.130	0.281	(-0.681,0.152)	6.61E-01	7.81	6	2.53E-01	-0.057	0.086	(0.111,-0.143)	5.07E-01	7.91	7	3.41E-01	-0.057	0.092	(0.123,-0.149)	5.35E-01	0.001	0.009	False	5.83E-02
Leukemia	LAM	30	0.059	0.132	(-0.2,0.192)	6.56E-01	21.25	28	8.15E-01	0.006	0.038	(0.08,-0.031)	8.68E-01	21.42	29	8.43E-01	0.017	0.040	(0.096,-0.023)	6.66E-01	0.002	0.030	False	3.55E-04
Lung	LAM	15	-0.211	0.259	(-0.72,0.048)	4.30E-01	8.62	13	8.01E-01	-0.092	0.076	(0.057,-0.169)	2.25E-01	8.85	14	8.41E-01	-0.092	0.083	(0.07,-0.174)	2.67E-01	0.001	0.014	False	1.96E-02
Melanoma	LAM	39	0.050	0.138	(-0.221,0.189)	7.18E-01	18.64	37	9.95E-01	0.042	0.059	(0.159,-0.017)	4.79E-01	18.64	38	9.96E-01	0.038	0.063	(0.162,-0.025)	5.44E-01	0.008	0.032	False	3.04E-03
Non-Hodukin lymphon	LAM	16	0.036	0.275	(-0.503,0.311)	8.98E-01	23.78	14	4.87E-02	-0.068	0.098	(0.125,-0.167)	4.88E-01	24.06	15	6.40E-02	-0.082	0.101	(0.115,-0.183)	4.13E-01	0.002	0.030	False	7.79E-05
Oral-pharynx	LAM	13	0.022	0.149	(-0.271,0.171)	8.87E-01	7.18	11	7.84E-01	-0.023	0.070	(0.114,-0.093)	7.37E-01	7.30	12	8.37E-01	-0.029	0.076	(0.12,-0.106)	6.99E-01	0.002	0.015	False	2.99E-02
Ovary	LAM	20	-0.073	0.107	(-0.282,0.034)	5.05E-01	11.53	18	8.76E-01	-0.062	0.052	(0.04,-0.114)	2.32E-01	11.54	19	9.04E-01	-0.067	0.056	(0.043,-0.123)	2.34E-01	0.006	0.019	False	1.00E-01
Pancreas	LAM	11	0.152	0.193	(-0.227,0.346)	4.51E-01	5.71	9	7.68E-01	0.055	0.056	(0.164,-0.001)	3.23E-01	5.99	10	8.16E-01	0.056	0.061	(0.175,-0.004)	3.54E-01	0.001	0.011	False	4.48E-02
PEF	LAM	189	-0.795	0.463	(-1.702,-0.332)	8.74E-02	168.29	187	8.33E-01	-0.211	0.159	(0.1,-0.369)	1.83E-01	170.10	188	8.21E-01	-0.239	0.165	(0.085,-0.404)	1.49E-01	0.036	0.166	False	2.11E-14
Prostate	LAM	76	0.193	0.122	(-0.046,0.315)	1.18E-01	62.51	74	8.27E-01	0.026	0.051	(0.125,-0.025)	6.10E-01	64.78	75	7.94E-01	0.026	0.053	(0.13,-0.027)	6.28E-01	0.022	0.085	False	2.31E-06
Rectum	LAM	24	0.055	0.149	(-0.237,0.204)	7.17E-01	11.69	22	9.63E-01	0.034	0.058	(0.147,-0.024)	5.60E-01	11.71	23	9.75E-01	0.032	0.063	(0.155,-0.031)	6.08E-01	0.003	0.020	False	7.95E-03
Thyroid	LAM	19	0.056	0.100	(-0.139,0.156)	5.79E-01	7.72	17	9.72E-01	-0.040	0.034	(0.027,-0.074)	2.46E-01	8.78	18	9.65E-01	-0.038	0.037	(0.035,-0.075)	3.06E-01	0.002	0.012	False	5.55E-02
LAM	Bladder	8	-0.156	0.215	(-0.577,0.059)	4.94E-01	5.98	6	4.26E-01	-0.165	0.057	(-0.277,-0.108)	3.75E-03	5.98	7	5.42E-01	-0.177	0.064	(-0.302,-0.113)	5.74E-03	0.184	3.76E-05	True	4.10E-54
LAM	Breast	8	0.082	0.095	(-0.104,0.177)	4.20E-01	9.01	6	1.73E-01	-0.009	0.026	(-0.06,0.016)	7.15E-01	10.52	7	1.61E-01	-0.008	0.027	(-0.062,0.019)	7.69E-01	0.184	4.57E-05	True	7.33E-54
LAM	Cervix	8	0.141	0.124	(-0.102,0.265)	2.99E-01	3.93	6	6.87E-01	0.001	0.033	(-0.064,0.034)	9.85E-01	5.31	7	6.73E-01	0.005	0.036	(-0.066,0.04)	8.99E-01	0.184	2.38E-05	True	2.78E-54
LAM	Colon	8	-0.036	0.241	(-0.509,0.266)	8.87E-01	12.73	6	4.75E-02	0.033	0.061	(-0.086,0.094)	5.87E-01	12.92	7	7.41E-02	0.026	0.060	(-0.091,0.086)	6.63E-01	0.184	3.66E-05	True	3.91E-54
LAM	Endometrial	8	-0.168	0.226	(-0.611,0.058)	4.86E-01	5.38	6	4.72E-01	-0.176	0.060	(-0.293,-0.116)	3.29E-03	5.38	7	5.90E-01	-0.186	0.067	(-0.317,-0.119)	5.54E-03	0.184	6.52E-05	True	1.53E-53
LAM	Esophagus	9	0.047	0.023	(0.002,0.07)	8.07E-02	23.53	7	1.38E-03	0.004	0.008	(-0.011,0.012)	5.89E-01	36.29	8	1.61E-05	0.006	0.006	(-0.006,0.012)	3.29E-01	0.184	9.59E-05	True	2.89E-53
LAM	FEV1	9	0.039	0.015	(0.01,0.054)	2.54E-02	9.37	7	2.27E-01	0.002	0.006	(-0.01,0.008)	7.71E-01	18.57	8	1.74E-02	0.001	0.006	(-0.01,0.006)	9.60E-01	0.184	4.80E-05	True	6.52E-54
LAM	FEV1/FVC	9	0.030	0.021	(-0.016,0.053)	2.43E-01	23.63	7	1.32E-03	0.003	0.007	(-0.011,0.01)	6.85E-01	28.53	8	3.83E-04	0.005	0.006	(-0.006,0.011)	3.57E-01	0.184	7.40E-05	True	1.54E-53
LAM	FVC	8	-0.424	0.314	(-1.038,-0.11)	2.25E-01	3.45	6	7.50E-01	0.032	0.082	(-0.128,0.114)	6.93E-01	5.72	7	5.73E-01	0.044	0.089	(-0.13,0.133)	6.17E-01	0.184	1.49E-05	True	1.22E-54
LAM	Kidney	8	-0.069	0.332	(-0.721,0.263)	8.42E-01	8.78	6	1.86E-01	-0.058	0.082	(-0.218,0.024)	4.83E-01	8.78	7	2.69E-01	-0.048	0.079	(-0.204,0.031)	5.45E-01	0.184	2.29E-05	True	2.00E-54
LAM	Leukemia	8	0.373	0.327	(-0.267,0.7)	2.97E-01	2.38	6	8.82E-01	0.152	0.090	(-0.025,0.242)	9.22E-02	2.88	7	8.96E-01	0.154	0.099	(-0.041,0.253)	1.22E-01	0.184	2.12E-05	True	1.81E-54
LAM	Lung	8	0.130	0.198	(-0.257,0.328)	5.34E-01	4.70	6	5.83E-01	-0.021	0.054	(-0.126,0.033)	6.95E-01	5.34	7	6.19E-01	-0.013	0.058	(-0.127,0.046)	8.27E-01	0.184	1.35E-05	True	1.10E-54
LAM	Melanoma	8	0.071	0.139	(-0.201,0.21)	6.26E-01	8.14	6	2.28E-01	0.036	0.034	(-0.031,0.069)	2.84E-01	8.23	7	3.13E-01	0.032	0.033	(-0.034,0.065)	3.46E-01	0.184	2.59E-05	True	2.34E-54
LAM	Non-Hodukin	8	0.014	0.252	(-0.48,0.266)	9.59E-01	9.47	6	1.49E-01	0.112	0.064	(-0.013,0.176)	7.99E-02	9.73	7	2.04E-01	0.118	0.062	(-0.004,0.181)	5.78E-02	0.184	3.70E-05	True	3.98E-54
LAM	Oral-pharynx	8	0.322	0.277	(-0.221,0.599)	2.89E-01	2.60	6	8.57E-01	-0.013	0.076	(-0.162,0.063)	8.67E-01	4.18	7	7.58E-01	-0.013	0.083	(-0.176,0.07)	8.74E-01	0.184	1.50E-05	True	1.30E-54
LAM	Ovary	8	-0.214	0.458	(-1.111,0.244)	6.57E-01	14.72	6	2.25E-02	-0.078	0.110	(-0.295,0.012)	4.77E-01	14.95	7	3.66E-02	-0.106	0.110	(-0.321,0.004)	3.33E-01	0.184	7.92E-05	True	2.17E-53
LAM	Pancreas	8	0.113	0.384	(-0.639,0.498)	7.78E-01	4.49	6	6.11E-01	0.061	0.103	(-0.141,0.164)	5.55E-01	4.51	7	7.20E-01	0.062	0.112	(-0.157,0.175)	5.78E-01	0.184	1.90E-05	True	1.59E-54
LAM	PEF	9	0.039	0.017	(0.006,0.056)	5.31E-02	10.61	7	1.56E-01	0.009	0.006	(-0.002,0.015)	1.00E-01	15.79	8	4.52E-02	0.007	0.005	(-0.004,0.013)	1.90E-01	0.184	6.05E-05	True	1.05E-53
LAM	Prostate	8	-0.047	0.101	(-0.246,0.054)	6.58E-01	3.85	6	6.98E-01	0.019	0.027	(-0.034,0.046)	4.88E-01	4.30	7	7.45E-01	0.019	0.030	(-0.039,0.049)	5.14E-01	0.184	2.56E-05	True	3.32E-54
LAM	Rectum	8	0.115	0.320	(-0.511,0.435)	7.31E-01	12.59	6	5.01E-02	-0.022	0.080	(-0.179,0.058)	7.80E-01	13.01	7	7.19E-02	-0.008	0.087	(-0.177,0.079)	9.28E-01	0.184	3.41E-05	True	3.51E-54
LAM	Thyroid	8	-0.260	0.373	(-0.992,0.113)	5.12E-01	6.23	6	3.98E-01	-0.012	0.097	(-0.202,0.085)	9.01E-01	6.73	7	4.58E-01	-0.026	0.105	(-0.231,0.079)	8.06E-01	0.184	1.72E-05	True	1.42E-54

Table S6. Gene candidates linked to lung function (Shrine et al., *Nat Genet* 2019) and coding for NR3C1 interactors.

Candidate gene name	Gene title	PubMed NR3C1 interaction
ALX1	ALX homeobox 1	https://pubmed.ncbi.nlm.nih.gov/31182584/
AP3B1	Adaptor related protein complex 3 beta 1 subunit	https://pubmed.ncbi.nlm.nih.gov/31182584/
CAMK2G	Calcium/calmodulin dependent protein kinase II gamma	https://pubmed.ncbi.nlm.nih.gov/31182584/
ELAVL2	ELAV like RNA binding protein 2	https://pubmed.ncbi.nlm.nih.gov/31182584/
FKBP4	FK506 binding protein 4	https://pubmed.ncbi.nlm.nih.gov/8341706/
JMJD1C	Jumonji domain containing 1C	https://pubmed.ncbi.nlm.nih.gov/28611094/
KIAA1462	Junctional cadherin 5 associated	https://pubmed.ncbi.nlm.nih.gov/31182584/
LMOD1	Leiomodin 1	https://pubmed.ncbi.nlm.nih.gov/31182584/
MTCL1	Microtubule crosslinking factor 1	https://pubmed.ncbi.nlm.nih.gov/31182584/
NCOR1	Nuclear receptor corepressor 1	https://pubmed.ncbi.nlm.nih.gov/12011091/
OTUD4	OTU deubiquitinase 4	https://pubmed.ncbi.nlm.nih.gov/31182584/
PITPNM3	PITPNM family member 3	https://pubmed.ncbi.nlm.nih.gov/31182584/
SATB2	SATB homeobox 2	https://pubmed.ncbi.nlm.nih.gov/31182584/
SMARCA2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A, member 2	https://pubmed.ncbi.nlm.nih.gov/17043312/
TRIP11	Thyroid hormone receptor interactor 11	https://pubmed.ncbi.nlm.nih.gov/31182584/
ZFP82	ZFP82 zinc finger protein	https://pubmed.ncbi.nlm.nih.gov/31182584/

Table S6. Gene candidates linked to lung function and coding for NR3C1 interactors.

References

1. Johnson SR, Cordier JF, Lazor R, Cottin V, Costabel U, Harari S, Reynaud-Gaubert M, Boehler A, Brauner M, Popper H, Bonetti F, Kingswood C, Review Panel of the ERS LAM Task Force. European Respiratory Society guidelines for the diagnosis and management of lymphangioleiomyomatosis. *Eur Respir J* 2010;35:14–26.

2. McCormack FX, Gupta N, Finlay GR, Young LR, Taveira-DaSilva AM, Glasgow CG, Steagall WK, Johnson SR, Sahn SA, Ryu JH, Strange C, Seyama K, Sullivan EJ, Kotloff RM, Downey GP, Chapman JT, Han MK, D'Armiento JM, Inoue Y, Henske EP, Bissler JJ, Colby TV, Kinder BW, Wikenheiser-Brokamp KA, Brown KK, Cordier JF, Meyer C, Cottin V, Brozek JL, *et al.* Official American Thoracic Society/Japanese Respiratory Society Clinical Practice guidelines: Lymphangioleiomyomatosis diagnosis and management. *Am J Respir Crit Care Med* 2016;194:748–761.

3. Martignoni G, Pea M, Reghellin D, Gobbo S, Zamboni G, Chilosi M, Bonetti F. Molecular pathology of lymphangioleiomyomatosis and other perivascular epithelioid cell tumors. *Arch Pathol Lab Med* 2010;134:33–40.

4. Harari S, Torre O, Cassandro R, Moss J. The changing face of a rare disease: Lymphangioleiomyomatosis. *Eur Respir J* 2015;46:1471–1485.

5. Dongre A, Clements D, Fisher AJ, Johnson SR. Cathepsin K in lymphangioleiomyomatosis: LAM cell-fibroblast interactions enhance protease activity by extracellular acidification. *Am J Pathol* 2017;187:1750–1762.

6. Henske EP, McCormack FX. Lymphangioleiomyomatosis - A wolf in sheep's clothing. *J Clin Invest* 2012;122:3807–3816.

7. Obraztsova K, Basil MC, Rue R, Sivakumar A, Lin SM, Mukhitov AR, Gritsiuta AI, Evans JF, Kopp M, Katzen J, Robichaud A, Atochina-Vasserman EN, Li S, Carl J, Babu A, Morley MP, Cantu E, Beers MF, Frank DB, Morrisey EE, Krymskaya VP. mTORC1 activation in lung mesenchyme drives sex- and age-dependent pulmonary structure and function decline. *Nat Commun* 2020;11:5640.

8. Guo M, Yu JJ, Perl AK, Wikenheiser-Brokamp KA, Riccetti M, Zhang EY, Sudha P, Adam M, Potter A, Kopras EJ, Giannikou K, Potter SS, Sherman S, Hammes SR, Kwiatkowski DJ, Whitsett JA, McCormack FX, Xu Y. Single cell transcriptomic analysis identifies a unique pulmonary lymphangioleiomyomatosis cell. *Am J Respir Crit Care Med* 2020;202:1373–1387.

9. Crino PB, Nathanson KL, Henske EP. The tuberous sclerosis complex. *N Engl J Med* 2006;355:1345–1356.

10. Giannikou K, Malinowska IA, Pugh TJ, Yan R, Tseng Y-Y, Oh C, Kim J, Tyburczy ME, Chekaluk Y, Liu Y, Alesi N, Finlay GA, Wu C-L, Signoretti S, Meyerson M, Getz G, Boehm JS, Henske EP, Kwiatkowski DJ. Whole exome sequencing identifies *TSC1/TSC2* biallelic loss as the primary and sufficient driver event for renal angiomyolipoma development. *PLoS Genet* 2016;12:e1006242.

11. Carsillo T, Astrinidis A, Henske EP. Mutations in the tuberous sclerosis complex gene *TSC2* are a cause of sporadic pulmonary lymphangioleiomyomatosis. *Proc Natl Acad Sci U S A* 2000;97:6085–6090.

12. McCormack FX, Inoue Y, Moss J, Singer LG, Strange C, Nakata K, Barker AF, Chapman JT, Brantly ML, Stocks JM, Brown KK, Lynch JP, Goldberg HJ, Young LR, Kinder BW, Downey GP, Sullivan EJ, Colby TV, McKay RT, Cohen MM, Korbee L, Taveira-DaSilva AM, Lee H-S, Krischer JP, Trapnell BC, National Institutes of Health Rare Lung Diseases Consortium, MILES Trial Group. Efficacy and safety of sirolimus in lymphangioleiomyomatosis. *N Engl J Med* 2011;364:1595–1606.

13. Krymskaya VP. Smooth muscle-like cells in pulmonary lymphangioleiomyomatosis. *Proc Am Thorac Soc* 2008;5:119–126.

14. Julian LM, Delaney SP, Wang Y, Goldberg AA, Doré C, Yockell-Lelièvre J, Tam RY, Giannikou K, McMurray F, Shoichet MS, Harper M-E, Henske EP, Kwiatkowski DJ, Darling TN, Moss J, Kristof AS, Stanford WL. Human pluripotent stem cell-derived TSC2-haploinsufficient smooth muscle cells recapitulate features of lymphangioleiomyomatosis. *Cancer Res* 2017;77:5491–5502.

15. Ruiz de Garibay G, Herranz C, Llorente A, Boni J, Serra-Musach J, Mateo F, Aguilar H, Gómez-Baldó L, Petit A, Vidal A, Climent F, Hernández-Losa J, Cordero Á, González-Suárez E, Sánchez-Mut JV, Esteller M, Llatjós R, Varela M, López JI, García N, Extremera AI, Gumà A, Ortega R, Plà MJ, Fernández A, Pernas S, Falo C, Morilla I, Campos M, *et al.* Lymphangioleiomyomatosis biomarkers linked to lung metastatic potential and cell stemness. *PLoS One* 2015;10:e0132546.

16. Pacheco-Rodríguez G, Steagall WK, Samsel L, Dagur PK, McCoy JP, Tunc I, Pirooznia M, Wang J-A, Darling TN, Moss J. Circulating lymphangioleiomyomatosis tumor cells with loss of heterozygosity in the *TSC2* gene show increased aldehyde dehydrogenase activity. *Chest* 2019;156:298–307.

17. McCormack FX, Travis WD, Colby TV, Henske EP, Moss J. Lymphangioleiomyomatosis: calling it what it is: A low-grade, destructive, metastasizing neoplasm. *Am J Respir Crit Care Med* 2012;186:1210–1212.

18. Liu H-J, Krymskaya VP, Henske EP. Immunotherapy for lymphangioleiomyomatosis and tuberous sclerosis: Progress and future directions. *Chest* 2019;156:1062–1067.

19. Krymskaya VP, McCormack FX. Lymphangioleiomyomatosis: A monogenic model of malignancy. *Annu Rev Med* 2017;68:69–83.

20. Kim W, Giannikou K, Dreier JR, Lee S, Tyburczy ME, Silverman EK, Radzikowska E, Wu S, Wu C-L, Henske EP, Hunninghake G, Carel H, Roman A, Pujana MA, Moss J, Won S, Kwiatkowski DJ. A genome-wide association study implicates NR2F2 in lymphangioleiomyomatosis pathogenesis. *Eur Respir J* 2019;53:.

21. Litchfield LM, Klinge CM. Multiple roles of COUP-TFII in cancer initiation and progression. *J Mol Endocrinol* 2012;49:R135–R148.

22. Polvani S, Pepe S, Milani S, Galli A. COUP-TFII in health and disease. *Cells* 2019;9:.

23. Rashkin SR, Graff RE, Kachuri L, Thai KK, Alexeeff SE, Blatchins MA, Cavazos TB, Corley DA, Emami NC, Hoffman JD, Jorgenson E, Kushi LH, Meyers TJ, Van Den Eeden SK, Ziv E, Habel LA, Hoffmann TJ, Sakoda LC, Witte JS. Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. *Nat Commun* 2020;11:4423.

24. Shrine N, Guyatt AL, Erzurumluoglu AM, Jackson VE, Hobbs BD, Melbourne CA, Batini C, Fawcett KA, Song K, Sakornsakolpat P, Li X, Boxall R, Reeve NF, Obeidat M, Zhao JH, Wielscher M, Weiss S, Kentistou KA, Cook JP, Sun BB, Zhou J, Hui J, Karrasch S, Imboden M, Harris SE, Marten J, Enroth S, Kerr SM, Surakka I, *et al.* New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nat Genet* 2019;51:481–493.

25. Rüeger S, McDaid A, Kutalik Z. Evaluation and application of summary statistic imputation to discover new height-associated loci. *PLoS Genet* 2018;14:e1007371.

26. Rüeger S, McDaid A, Kutalik Z. Improved imputation of summary statistics for admixed populations. *BioRxiv* 2018; doi.org/10.1101/203927.

27. Smeland OB, Bahrami S, Frei O, Shadrin A, O'Connell K, Savage J, Watanabe K, Krull F, Bettella F, Steen NE, Ueland T, Posthuma D, Djurovic S, Dale AM, Andreassen OA. Genome-wide analysis reveals extensive genetic overlap between schizophrenia, bipolar disorder, and intelligence. *Mol Psychiatry* 2020;25:844–853.

28. Ning Z, Pawitan Y, Shen X. High-definition likelihood inference of genetic correlations across human complex traits. *Nat Genet* 2020;52:859–864.

29. Andreassen OA, Thompson WK, Schork AJ, Ripke S, Mattingsdal M, Kelsoe JR, Kendler KS, O'Donovan MC, Rujescu D, Werge T, Sklar P, Psychiatric Genomics Consortium (PGC), Bipolar Disorder and Schizophrenia Working Groups, Roddey JC, Chen C-H, McEvoy L, Desikan RS, Djurovic S, Dale AM. Improved detection of common variants associated with schizophrenia and bipolar disorder using pleiotropy-informed conditional false discovery rate. *PLoS Genet* 2013;9:e1003455.

30. Liu JZ, Hov JR, Folseraas T, Ellinghaus E, Rushbrook SM, Doncheva NT, Andreassen OA, Weersma RK, Weismüller TJ, Eksteen B, Invernizzi P, Hirschfield GM, Gotthardt DN, Pares A, Ellinghaus D, Shah T, Juran BD, Milkiewicz P, Rust C, Schramm C, Müller T, Srivastava B, Dalekos G, Nöthen MM, Herms S, Winkelmann J, Mitrovic M, Braun F, Ponsioen CY, *et al.* Dense genotyping of immune-related disease regions identifies nine new risk loci for primary sclerosing cholangitis. *Nat Genet* 2013;45:670–675.

31. Schork AJ, Wang Y, Thompson WK, Dale AM, Andreassen OA. New statistical approaches exploit the polygenic architecture of schizophrenia -- Implications for the underlying neurobiology. *Curr Opin Neurobiol* 2016;36:89–98.

32. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, Hartwig FP, Holmes MV, Minelli C, Relton CL, Theodoratou E. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res* 2019;4:186.

33. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;44:512–525.

34. Zhao Q, Wang J, Hemani G, Bowden J, Small DS. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. *ArXiv180109652 Math Stat* 2019; arXiv:1801.09652.

35. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–560.

36. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet* 2017;13:e1007081.

37. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R, Tan VY, Yarmolinsky J, Shihab HA, Timpson NJ, Evans DM, Relton C, Martin RM, Davey Smith G, Gaunt TR, Haycock PC. The MR-Base platform supports systematic causal inference across the human phenome. *eLife* 2018;7.

38. Melé M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, Young TR, Goldmann JM, Pervouchine DD, Sullivan TJ, Johnson R, Segrè AV, Djebali S, Niarchou A, GTEx Consortium, Wright FA, Lappalainen T, Calvo M, Getz G, Dermitzakis ET, Ardlie KG, Guigó R. Human genomics. The human transcriptome across tissues and individuals. *Science* 2015;348:660–665.

39. Wang Y, Song F, Zhang B, Zhang L, Xu J, Kuang D, Li D, Choudhary MNK, Li Y, Hu M, Hardison R, Wang T, Yue F. The 3D Genome Browser: A web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *Genome Biol* 2018;19:151.

40. Luck K, Kim D-K, Lambourne L, Spirohn K, Begg BE, Bian W, Brignall R, Cafarelli T, Campos-Laborie FJ, Charloteaux B, Choi D, Coté AG, Daley M, Deimling S, Desbuleux A, Dricot A, Gebbia M, Hardy MF, Kishore N, Knapp JJ, Kovács IA, Lemmens I, Mee MW, Mellor JC, Pollis C, Pons C, Richardson AD, Schlabach S, Teeking B, *et al.* A reference map of the human binary protein interactome. *Nature* 2020;580:402–408.

41. Oughtred R, Rust J, Chang C, Breitkreutz B-J, Stark C, Willems A, Boucher L, Leung G, Kolas N, Zhang F, Dolma S, Coulombe-Huntington J, Chatr-Aryamontri A, Dolinski K, Tyers M. The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci* 2020;30:187–200.

42. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018;36:411–420.

43. Martin KR, Zhou W, Bowman MJ, Shih J, Au KS, Dittenhafer-Reed KE, Sisson KA, Koeman J, Weisenberger DJ, Cottingham SL, DeRoos ST, Devinsky O, Winn ME, Cherniack AD, Shen H, Northrup H, Krueger DA, MacKeigan JP. The genomic landscape of tuberous sclerosis complex. *Nat Commun* 2017;8:15816.

44. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, Kovatich AJ, Benz CC, Levine DA, Lee AV, Omberg L, Wolf DM, Shriver CD, Thorsson V, Cancer Genome Atlas Research Network, Hu H. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* 2018;173:400-416.e11.

45. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012;2:401–404.

46. Travaglini KJ, Nabhan AN, Penland L, Sinha R, Gillich A, Sit RV, Chang S, Conley SD, Mori Y, Seita J, Berry GJ, Shrager JB, Metzger RJ, Kuo CS, Neff N, Weissman IL, Quake SR, Krasnow MA. A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* 2020;587:619–625.

47. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinforma* 2016;32:2847–2849.

48. Nuñez O, Román A, Johnson SR, Inoue Y, Hirose M, Casanova Á, de Garibay GR, Herranz C, Bueno-Moreno G, Boni J, Mateo F, Petit A, Climent F, Soler T, Vidal A, Sánchez-Mut JV, Esteller M, López JI, García N, Gumà A, Ortega R, Plà MJ, Campos M, Ansótegui E, Molina-Molina M, Valenzuela C, Ussetti P, Laporta R, Ancochea J, *et al.* Study of breast cancer incidence in patients of lymphangioleiomyomatosis. *Breast Cancer Res Treat* 2016;156:195–201.

49. Nuñez O, Baldi BG, Radzikowska E, Carvalho CRR, Herranz C, Sobiecka M, Torre O, Harari S, Vergeer MAMH, Kolbe J, Pollán M, Pujana MA. Risk of breast cancer in patients with lymphangioleiomyomatosis. *Cancer Epidemiol* 2019;61:154–156.

50. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, McMahon A, Morales J, Mountjoy E, Sollis E, Suveges D, Vrousgou O, Whetzel PL, Amode R, Guillen JA, Riat HS, Trevanion SJ, Hall P, Junkins H, Flicek P, Burdett T, Hindorff LA, Cunningham F, Parkinson H. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 2019;47:D1005–D1012.

51. Beck T, Hastings RK, Gollapudi S, Free RC, Brookes AJ. GWAS Central: A comprehensive resource for the comparison and interrogation of genome-wide association studies. *Eur J Hum Genet* 2014;22:949–952.

52. Shaffer JR, Li J, Lee MK, Roosenboom J, Orlova E, Adhikari K, 23andMe Research Team, Gallo C, Poletti G, Schuler-Faccini L, Bortolini M-C, Canizales-Quinteros S, Rothhammer F, Bedoya G, González-José R, Pfeffer PE, Wollenschlaeger CA, Hecht JT, Wehby GL, Moreno LM, Ding A, Jin L, Yang Y, Carlson JC, Leslie EJ, Feingold E, Marazita ML, Hinds DA, Cox TC, *et al.* Multiethnic GWAS reveals polygenic architecture of earlobe attachment. *Am J Hum Genet* 2017;101:913–924.

53. Michailidou K, Beesley J, Lindstrom S, Canisius S, Dennis J, Lush MJ, Maranian MJ, Bolla MK, Wang Q, Shah M, Perkins BJ, Czene K, Eriksson M, Darabi H, Brand JS, Bojesen SE, Nordestgaard BG, Flyger H, Nielsen SF, Rahman N, Turnbull C, Bocs, Fletcher O, Peto J, Gibson L, dos-Santos-Silva I, Chang-Claude J, Flesch-Janys D, Rudolph A, *et al.* Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat Genet* 2015;47:373–380.

54. Liyanage UE, Law MH, Han X, An J, Ong J-S, Gharahkhani P, Gordon S, Neale RE, Olsen CM, 23andMe Research Team, MacGregor S, Whiteman DC. Combined analysis of keratinocyte cancers identifies novel genome-wide loci. *Hum Mol Genet* 2019;28:3148–3160.

55. COGENT Study, Houlston RS, Webb E, Broderick P, Pittman AM, Di Bernardo MC, Lubbe S, Chandler I, Vijayakrishnan J, Sullivan K, Penegar S, Colorectal Cancer Association Study Consortium, Carvajal-Carmona L, Howarth K, Jaeger E, Spain SL, Walther A, Barclay E, Martin L, Gorman M, Domingo E, Teixeira AS, CoRGI Consortium, Kerr D, Cazier J-B, Niittymäki I, Tuupanen S, Karhu A, Aaltonen LA, *et al.* Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. *Nat Genet* 2008;40:1426–1435.

56. Jeong C, Witonsky DB, Basnyat B, Neupane M, Beall CM, Childs G, Craig SR, Novembre J, Di Rienzo A. Detecting past and ongoing natural selection among ethnically Tibetan women at high altitude in Nepal. *PLoS Genet* 2018;14:e1007650.

57. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, Mead D, Bouman H, Riveros-Mckay F, Kostadima MA, Lambourne JJ, Sivapalaratnam S, Downes K, Kundu K, Bomba L, Berentsen K, Bradley JR, Daugherty LC, Delaneau O, Freson K, Garner SF, Grassi L, Guerrero J, Haimel M, Janssen-Megens EM, Kaan A, Kamat M, Kim B, Mandoli A, *et al.* The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell* 2016;167:1415-1429.e19.

58. Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D, Werge T, Pietiläinen OPH, Mors O, Mortensen PB, Sigurdsson E, Gustafsson O, Nyegaard M, Tuulio-Henriksson A, Ingason A, Hansen T, Suvisaari J, Lonnqvist J, Paunio T, Børglum AD, Hartmann A, Fink-Jensen A, Nordentoft M, Hougaard D, Norgaard-Pedersen B, Böttcher Y, Olesen J, Breuer R, Möller H-J, *et al.* Common variants conferring risk of schizophrenia. *Nature* 2009;460:744–747.

59. De Martino MU, Bhattachryya N, Alesci S, Ichijo T, Chrousos GP, Kino T. The glucocorticoid receptor and the orphan nuclear receptor chicken ovalbumin upstream promoter-transcription factor II interact with and mutually affect each other's transcriptional activities: Implications for intermediary metabolism. *Mol Endocrinol* 2004;18:820–833.

60. Regan Anderson TM, Ma SH, Raj GV, Cidlowski JA, Helle TM, Knutson TP, Krutilina RI, Seagroves TN, Lange CA. Breast tumor kinase (Brk/PTK6) is induced by HIF, glucocorticoid receptor, and PELP1-mediated stress signaling in triple-negative breast cancer. *Cancer Res* 2016;76:1653–1663.

61. Biadasiewicz K, Fock V, Dekan S, Proestling K, Velicky P, Haider S, Knöfler M, Fröhlich C, Pollheimer J. Extravillous trophoblast-associated ADAM12 exerts pro-invasive properties, including induction of integrin beta 1-mediated cellular spreading. *Biol Reprod* 2014;90:101.

62. Pittman AM, Naranjo S, Jalava SE, Twiss P, Ma Y, Olver B, Lloyd A, Vijayakrishnan J, Qureshi M, Broderick P, van Wezel T, Morreau H, Tuupanen S, Aaltonen LA, Alonso ME, Manzanares M, Gavilán A, Visakorpi T, Gómez-Skarmeta JL, Houlston RS. Allelic variation at the 8q23.3 colorectal cancer risk locus functions as a cis-acting regulator of *EIF3H*. *PLoS Genet* 2010;6:e1001126.

63. Chen Z, Lan X, Wu D, Sunkel B, Ye Z, Huang J, Liu Z, Clinton SK, Jin VX, Wang Q. Ligand-dependent genomic function of glucocorticoid receptor in triple-negative breast cancer. *Nat Commun* 2015;6:8323.

64. Sorrentino G, Ruggeri N, Zannini A, Ingallina E, Bertolio R, Marotta C, Neri C, Cappuzzello E, Forcato M, Rosato A, Mano M, Bicciato S, Del Sal G. Glucocorticoid receptor signalling activates YAP in breast cancer. *Nat Commun* 2017;8:14073.

65. Obradović MMS, Hamelin B, Manevski N, Couto JP, Sethi A, Coissieux M-M, Münst S, Okamoto R, Kohler H, Schmidt A, Bentires-Alj M. Glucocorticoids promote breast cancer metastasis. *Nature* 2019;567:540–544.

66. West DC, Kocherginsky M, Tonsing-Carter EY, Dolcen DN, Hosfield DJ, Lastra RR, Sinnwell JP, Thompson KJ, Bowie KR, Harkless RV, Skor MN, Pierce CF, Styke SC, Kim CR, de Wet L, Greene GL, Boughey JC, Goetz MP, Kalari KR, Wang L, Fleming GF, Györffy B, Conzen SD. Discovery of a glucocorticoid receptor (GR) activity signature using selective GR antagonism in ER-negative breast cancer. *Clin Cancer Res* 018;24:3433–3446.

67. Beesley J, Sivakumaran H, Moradi Marjaneh M, Shi W, Hillman KM, Kaufmann S, Hussein N, Kar S, Lima LG, Ham S, Möller A, Chenevix-Trench G, Edwards SL, French JD. eQTL colocalization analyses identify *NTN4* as a candidate breast cancer risk gene. *Am J Hum Genet* 2020;107:778–787.

68. Kociok N, Crespo-Garcia S, Liang Y, Klein SV, Nürnberg C, Reichhart N, Skosyrski S, Moritz E, Maier A-K, Brunken WJ, Strauß O, Koch M, Joussen AM. Lack of netrin-4 modulates pathologic neovascularization in the eye. *Sci Rep* 2016;6:18828.

69. Xu X, Yan Q, Wang Y, Dong X. NTN4 is associated with breast cancer metastasis via regulation of EMT-related biomarkers. *Oncol Rep* 2017;37:449–457.

70. Derfuss T, Parikh K, Velhin S, Braun M, Mathey E, Krumbholz M, Kümpfel T, Moldenhauer A, Rader C, Sonderegger P, Pöllmann W, Tiefenthaller C, Bauer J, Lassmann H, Wekerle H, Karagogeos D, Hohlfeld R, Linington C, Meinl E. Contactin-2/TAG-1-directed autoimmunity is identified in multiple sclerosis patients and mediates gray matter pathology in animals. *Proc Natl Acad Sci U S A* 2009;106:8302–8307.

71. Abou-Kandil A, Eisa N, Jabareen A, Huleihel M. Differential effects of HTLV-1 Tax oncoprotein on the different estrogen-induced-ER α-mediated transcriptional activities. *Cell Cycle* 2016;15:2626–2635.

72. Yin GN, Lee HW, Cho J-Y, Suk K. Neuronal pentraxin receptor in cerebrospinal fluid as a potential biomarker for neurodegenerative diseases. *Brain Res* 2009;1265:158–170.

73. Smith GD, Ebrahim S. “Mendelian randomization”: Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;32:1–22.

74. Delaney SP, Julian LM, Stanford WL. The neural crest lineage as a driver of disease heterogeneity in Tuberous Sclerosis Complex and Lymphangioleiomyomatosis. *Front Cell Dev Biol* 2014;2:69.

75. Yue M, Pacheco G, Cheng T, Li J, Wang Y, Henske EP, Schuger L. Evidence supporting a lymphatic endothelium origin for angiomyolipoma, a TSC2-tumor related to lymphangioleiomyomatosis. *Am J Pathol* 2016;186:1825–1836.

76. Hayashi T, Kumasaka T, Mitani K, Terao Y, Watanabe M, Oide T, Nakatani Y, Hebisawa A, Konno R, Takahashi K, Yao T, Seyama K. Prevalence of uterine and adnexal involvement in pulmonary lymphangioleiomyomatosis: A clinicopathologic study of 10 patients. *Am J Surg Pathol* 2011;35:1776–1785.

77. Prizant H, Sen A, Light A, Cho S-N, DeMayo FJ, Lydon JP, Hammes SR. Uterine-specific loss of Tsc2 leads to myometrial tumors in both the uterus and lungs. *Mol Endocrinol* 2013;27:1403–1414.

78. Zhao F, Franco HL, Rodriguez KF, Brown PR, Tsai M-J, Tsai SY, Yao HH-C. Elimination of the male reproductive tract in the female embryo is promoted by COUP-TFII in mice. *Science* 2017;357:717–720.

79. Dwyer AR, Truong TH, Ostrander JH, Lange CA. 90 YEARS OF PROGESTERONE: Steroid receptors as MAPK signaling sensors in breast cancer: let the fates decide. *J Mol Endocrinol* 2020;65:T35–T48.

80. Prizant H, Hammes SR. Minireview: Lymphangioleiomyomatosis (LAM): The “Other” Steroid-Sensitive Cancer. *Endocrinology* 2016;157:3374–3383.

81. Pan D, Kocherginsky M, Conzen SD. Activation of the glucocorticoid receptor is associated with poor prognosis in estrogen receptor-negative breast cancer. *Cancer Res* 2011;71:6360–6370.

82. Wu F, Chen W, Kang X, Jin L, Bai J, Zhang H, Zhang X. A seven-nuclear receptor-based prognostic signature in breast cancer. *Clin Transl Oncol* 2020;doi:10.1007/s12094-020-02517-1.

83. Kach J, Conzen SD, Szmulewitz RZ. Glucocorticoid receptor signaling in breast and prostate cancers: Emergence as a therapeutic target. *Sci Transl Med* 2015;7:305ps19.

84. Sun Y, Zhang E, Lao T, Pereira AM, Li C, Xiong L, Morrison T, Haley KJ, Zhou X, Yu JJ. Progesterone and estradiol synergistically promote the lung metastasis of tuberin-deficient cells in a preclinical model of lymphangioleiomyomatosis. *Horm Cancer* 2014;5:284–298.

85. Li X, Liu X, Zhang L, Li C, Zhang E, Ma W, Fan Q, Yu JJ. Insulin growth factor binding protein 2 mediates the progression of lymphangioleiomyomatosis. *Oncotarget* 2017;8:36628–36638.

86. Valencia JC, Matsui K, Bondy C, Zhou J, Rasmussen A, Cullen K, Yu ZX, Moss J, Ferrans VJ. Distribution and mRNA expression of insulin-like growth factor system in pulmonary lymphangioleiomyomatosis. *J Investig Med* 2001;49:421–433.

87. Matsui K, Takeda K, Yu ZX, Travis WD, Moss J, Ferrans VJ. Role for activation of matrix metalloproteinases in the pathogenesis of pulmonary lymphangioleiomyomatosis. *Arch Pathol Lab Med* 2000;124:267–275.

88. Glassberg MK, Elliot SJ, Fritz J, Catanuto P, Potier M, Donahue R, Stetler-Stevenson W, Karl M. Activation of the estrogen receptor contributes to the progression of pulmonary lymphangioleiomyomatosis via matrix metalloproteinase-induced cell invasiveness. *J Clin Endocrinol Metab* 2008;93:1625–1633.

89. Goncharova EA, Goncharov DA, Fehrenbach M, Khavin I, Ducka B, Hino O, Colby TV, Merrilees MJ, Haczku A, Albelda SM, Krymskaya VP. Prevention of alveolar destruction and airspace enlargement in a mouse model of pulmonary lymphangioleiomyomatosis (LAM). *Sci Transl Med* 2012;4:154ra134.

90. Huo Y, Li S, Liu J, Li X, Luo X-J. Functional genomics reveal gene regulatory mechanisms underlying schizophrenia risk. *Nat Commun* 2019;10:670.

91. Middleton EA, Weyrich AS, Zimmerman GA. Platelets in pulmonary immune responses and inflammatory lung diseases. *Physiol Rev* 2016;96:1211–1259.

92. Bridges JP, Sudha P, Lipps D, Wagner A, Guo M, Du Y, Brown K, Filuta A, Kitzmiller J, Stockman C, Chen X, Weirauch MT, Jobe AH, Whitsett JA, Xu Y. Glucocorticoid regulates mesenchymal cell differentiation required for perinatal lung morphogenesis and function. *Am J Physiol Lung Cell Mol Physiol* 2020;319:L239–L255.

93. Miller S, Stewart ID, Clements D, Soomro I, Babaei-Jadidi R, Johnson SR. Evolution of lung pathology in lymphangioleiomyomatosis: Associations with disease course and treatment response. *J Pathol Clin Res* 2020;6:215–226.

94. Espín R, Baiges A, Blommaert E, Herranz C, Roman A, Saez B, Ancochea J, Valenzuela C, Ussetti P, Laporta R, Rodriguez-Portal JA, van Moorsel CHM, van der Vis JJ, Quanjel MJR, Villar-Piqué A, Diaz-Lucena D, Llorens F, Casanova A, Molina-Molina M, Plass M, Mateo F, Moss J, Pujana MA. Heterogeneity and cancer-related features in lymphangioleiomyomatosis cells and tissue. *Mol Cancer Res* 2021; doi:10.1158/1541-7786.MCR-21-0220.

95. Hardavella G, Tzortzaki EG, Siozopoulou V, Galanis P, Vlachaki E, Avgousti M, Stefanou D, Siafakas NM. Lymphangiogenesis in COPD: Another link in the pathogenesis of the disease. *Respir Med* 2012;106:687–693.

96. Schulz M, Eggert M, Baniahmad A, Dostert A, Heinzel T, Renkawitz R. RU486-induced glucocorticoid receptor agonism is controlled by the receptor N terminus and by corepressor binding. *J Biol Chem* 2002;277:26238–26243.