Genome-Wide Association Study of Susceptibility to Idiopathic Pulmonary Fibrosis

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Genome-Wide Association Study of Susceptibility to Idiopathic Pulmonary Fibrosis

Richard J Allen[1], Beatriz Guillen-Guio[2], Justin M Oldham[3], Shwu-Fan Ma[4], Amy Dressen[5], Megan L Paynton[1], Luke M Kraven[1], Ma'en Obeidat[6], Xuan Li[6], Michael Ng[7], Rebecca Braybrooke[8,9], Maria Molina-Molina[10,11,12], Brian D Hobbs[13,14], Rachel K Putman[14], Phuwanat Sakornsakolpat[13,15], Helen L Booth[16], William A Fahy[17], Simon P Hart[18], Mike R Hill[19], Nik Hirani[20], Richard B Hubbard[8,9], Robin J McAnulty[21], Ann B Millar[22], Vidyia Navaratnam[8,9], Eunice Oballa[17], Helen Parfrey[23], Gauri Saini[24], Moira K B Whyte[20], Yingze Zhang[25,26], Naftali Kaminski[27], Ayodeji Adegunsoye[28], Mary E Strek[28], Margaret Neighbors[5], Xuting R Sheng[5], Gunnar Gudmundsson[29,30], Vilmundur Gudnason[31,32], Hiroto Hatabu[33,34], David J Lederer[35,36], Ani Manichaikul[37,38], John D Newell, Jr[39,40], George T O'Connor[41,42], Victor E Ortega[43], Hanfei Xu[44], Tasha E Fingerlin[45,46], Yohan Bossé[47], Ke Hao[48,49], Philippe Joubert[47], David C Nickle[50], Don D Sin[6,51], Wim Timens[52,53], Dominic Furniss[7], Andrew P Morris[54,55,56], Krina Zondervan[54], Ian P Hall[9,57], Ian Sayers[9,57], Martin D Tobin[1,58], Toby M Maher[59,60], Michael H Cho[13,14], Gary M Hunninghake[14,34], David A Schwartz[45,61,62], Brian L Yaspan[5], Philip L Molyneaux[59,60], Carlos Flores[2,12,63,64], Imre Noth[4], R Gisli Jenkins[9,57]*, Louise V Wain[1,58]*

[1]Department of Health Sciences, University of Leicester, Leicester, UK,

[2]Research Unit, Hospital Universitario N.S. de Candelaria, Universidad de La Laguna, Santa Cruz de Tenerife, Spain,

[3]Department of Internal Medicine, University of California Davis, Davis, USA,

[4]Division of Pulmonary & Critical Care Medicine, University of Virginia, Charlottesville, USA,

[5]Genentech, South San Francisco, USA,

[6]The University of British Columbia Centre for Heart Lung Innovation, St Paul's Hospital,

Vancouver, Canada,

[7]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK,

[8]Division of Epidemiology and Public Health, University of Nottingham, Nottingham, UK,

[9]National Institute for Health Research, Nottingham Biomedical Research Centre, Nottingham University Hospitals, Nottingham, UK,

[10]Servei de Pneumologia, Laboratori de Pneumologia Experimental, IDIBELL, Spain

[11]Campus de Bellvitge, Universitat de Barcelona, Barcelona, Spain,

[12]CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Spain,

[13]Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, USA,

[14]Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, USA,

[15]Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand,

[16]Department of Thoracic Medicine, University College London Hospitals, London, UK,

[17]Discovery Medicine, GlaxoSmithKline, Stevenage, UK,

[18]Respiratory Research Group, Hull York Medical School, Castle Hill Hospital, Cottingham, UK,

[19]Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, Oxford, UK,

[20]MRC Centre for Inflammation Research at the University of Edinburgh, Edinburgh, UK,

[21]UCL Respiratory Centre for Inflammation and Tissue Repair, University College London, London, UK,

[22]Academic Respiratory Unit, School of Clinical Sciences, University of Bristol, Bristol, UK,

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

[23]Cambridge ILD Service, Royal Papworth Hospital, Cambridge, UK,

[24]Respiratory Medicine, Nottingham University Hospitals Trust, UK,

[25]Division of Pulmonary, Allergy and Critical Care Medicine, University of Pittsburgh, Pittsburgh, USA,

[26]Simmons Center for Interstitial Lung Diseases, University of Pittsburgh, Pittsburgh, USA,

[27]Section of Pulmonary , Critical Care and Sleep Medicine, Yale School of Medicine, New Haven, USA,

[28]Section of Pulmonary & Critical Care, Department of Medicine; The University of Chicago, Chicago, USA,

[29]Dept of Respiratory Medicine, Landspital University Hospital, Reykjavik, Iceland,

[30]Faculty of Medicine, University of Iceland, Reykjavik, Iceland,

[31]University of Iceland, Reykjavik, Iceland,

[32]Icelandic Heart Association, Kopavogur, Iceland,

[33]Department of Radiology, Brigham and Women's Hospital, Boston, USA,

[34]Center for Pulmonary Functional Imaging, Brigham and Women's Hospital, Boston, USA,

[35]Department of Medicine, College of Physicians and Surgeons, Columbia University, New York, USA,

[36]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, USA,

[37]Center for Public Health Genomics, University of Virginia, Charlottesville, USA,

[38]Department of Public Health Sciences, University of Virginia, Charlottesville, USA,

[39]Division of Cardiovascular and Pulmonary Imaging, Department of Radiology, University of Iowa Carver College of Medicine, Iowa City, USA,

[40]Department of Radiology, University of Washington, Seattle, USA,

[41]Pulmonary Center, Department of Medicine, Boston University, Boston, USA,

[42]National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham,USA,

[43]Center for Genomics and Personalized Medicine, Wake Forest School of Medicine, Winston-Salem, USA,

[44]Department of Biostatistics, Boston University School of Public Health, Boston, USA,

[45]Center for Genes, Environment and Health, National Jewish Health, Denver, USA,

[46]Department of Biostatistics and Informatics, University of Colorado, Denver, USA,

[4743]Institut universitaire de cardiologie et de pneumologie de Québec, Université Laval, Quebec City, Canada,

[48]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, USA,

[49]Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, USA,

[50]Merck Research Laboratories, Genetics and Pharmacogenomics, Boston, USA,

[51]Respiratory Division, Department of Medicine, University of British Columbia, Vancouver, Canada,

[52]University of Groningen, University Medical Center Groningen, Department of Pathology and Medical Biology, Groningen, The Netherlands,

[53]University of Groningen, University Medical Center Groningen, Groningen Research Institute for Asthma and COPD, Groningen, The Netherlands,

[54]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK,

[55]Department of Biostatistics, University of Liverpool, Liverpool, UK,

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

[56]Division of Musculoskeletal and Dermatological Sciences, University of Manchester, Manchester, UK

[57]Division of Respiratory Medicine, University of Nottingham, Nottingham, UK,

[58]National Institute for Health Research, Leicester Respiratory Biomedical Research Centre, Glenfield Hospital, Leicester, UK,

[59]NIHR Respiratory Clinical Research Facility, Royal Brompton Hospital, London, UK,

[60]National Heart and Lung Institute, Imperial College, London, UK,

[61]Department of Medicine, University of Colorado Denver, Denver, USA,

[62]Department of Immunology, University of Colorado Denver, Denver, USA,

[63]Instituto Tecnológico y de Energías Renovables (ITER, S.A.), Santa Cruz de Tenerife, Spain,

[64]Instituto de Tecnologías Biomédicas (ITB), Universidad de La Laguna, Santa Cruz de Tenerife, Spain,

* Contributed equally

Corresponding author: Louise Wain (email: lvw1@leicester.ac.uk, phone: 0116 229 7252)

Author contributions

LVW, RGJ, IN, CF, RJA and JMO designed the study. RJA, BGG, AD, BLY, S-FM, MiN, MLP, LMK, MO, XL, BDH, RKP and PS analysed the data. RGJ, LVW, IN, DAS, TEF, CF, JMO, SFM, RB, MMM, HLB, WAF, SPH, MRH, NH, RBH, RJM, ABM, VN, EO, HP, GS, MKBW, YZ, NK, AA, MES, MaN, XS, IPH, IS, MDT, TMM, BLY, PLM MDH, RKP, PS, GG, VG, HH, DLJ, AM, JDN, GTO'C, VEO, HX, MHC, GMH, MO, YB, KH, PJ, DCN, DDS and WT were responsible for recruitment, screening and genotyping of cases and controls for IPF, ILA and gene expression analyses. LVW, RGJ, IN, CF, JMO and DAS supervised and coordinated the study. RJA, RGJ and LVW led

the writing of the manuscript. All authors contributed to drafting and providing critical feedback on the manuscript.

Funding

R. Allen is an Action for Pulmonary Fibrosis Research Fellow. L. Wain holds a GSK/British Lung Foundation Chair in Respiratory Research. RG. Jenkins is supported by an NIHR Research Professorship (NIHR reference RP-2017-08-ST2-014). I. Noth: National Heart Lung and Blood Institute (R01HL130796). B. Guillen is funded by Agencia Canaria de Investigación, Innovación y Sociedad de la Información (TESIS2015010057) co-funded by European Social Fund. J. Oldham: National Heart Lung and Blood Institute (K23HL138190). C. Flores: Spanish Ministry of Science, Innovation and Universities (grant RTC-2017-6471-1; MINECO/AEI/FEDER, UE) co-financed by the European Regional Development Funds (ERDF) 'A way of making Europe' from the European Union, and by the agreement OA17/008 with Instituto Tecnológico y de Energías Renovables (ITER) to strengthen scientific and technological education, training, research, development and innovation in Genomics, Personalized Medicine and Biotechnology. The Spain Biobank array genotyping service was carried out at CEGEN-PRB3-ISCIII; which is supported by PT17/0019, of the PE I+D+i 2013-2016, funded by Instituto de Salud Carlos III, and co-financed by ERDF. P. Molyneaux is an Action for Pulmonary Fibrosis Research Fellow. M. Obeidat is a fellow of the Parker B Francis Foundation and a Scholar of the Michael Smith Foundation for Health Research (MSFHR). B. Hobbs: NIH K08 HL136928, Parker B. Francis Research Opportunity Award. M. Cho and G Hunninghake: This work was supported by NHLBI grants R01HL113264 (M.H.C), R01HL137927 (M.H.C.), R01HL135142 (M.H.C. and G.M.H) and R01111024 (G.M.H.). The content is solely the responsibility of the authors and does not necessarily represent the

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

official views of the NIH. The funding body has no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript. T. Maher is supported by an NIHR Clinician Scientist Fellowship (NIHR Ref: CS-2013-13-017) and a British Lung Foundation Chair in Respiratory Research (C17-3). M. Tobin is supported by a Wellcome Trust Investigator Award (WT202849/Z/16/Z). The research was partially supported by the National Institute for Health Research (NIHR) Leicester Biomedical Research Centre; the views expressed are those of the author(s) and not necessarily those of the National Health Service (NHS), the NIHR or the Department of Health. I. Hall was partially supported by the NIHR Nottingham Biomedical Research Centre; the views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. I. Sayers: MRC (G1000861) and Asthma UK (AUK-PG-2013-188). D. Furniss was supported by an Intermediate Fellowship from the Wellcome Trust (097152/Z/11/Z). This work was partially supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre. V. Navaratnam is funded by an NIHR Clinical Lectureship. G. Gudmundsson is supported by project grant 141513-051 from the Icelandic Research Fund and Landspitali Scientific Fund A-2016-023, A-2017-029 and A-2018-025. A. Manichaikul and D. Lederer: MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-001079, UL1-TR-001420, UL1-TR-001881, and DK063491. Funding for SHARe genotyping was provided by NHLBI Contract N02-HL-64278. Genotyping was performed at Affymetrix (Santa Clara, California, USA) and the Broad Institute of Harvard

and MIT (Boston, Massachusetts, USA) using the Affymetrix Genome-Wide Human SNP Array 6.0. This work was supported by NIH grants R01 HL131565 (A.M.), R01 HL103676 (D.J.L.) and R01 HL137234 (D.J.L.).

Shortened title: Genome-wide study of idiopathic pulmonary fibrosis

Subject category: 9.23 Interstitial Lung Disease

Word count: 3,670

Some of the results of these studies have been previously reported in the form of a preprint (bioRxiv, 14 May 2019 https://www.biorxiv.org/content/10.1101/636761v1).

This article has an online data supplement, which is accessible from this issue's table of content online at www.atsjournals.org.

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Abstract

Rationale: Idiopathic pulmonary fibrosis (IPF) is a complex lung disease characterised by scarring of the lung that is believed to result from an atypical response to injury of the epithelium. Genome-wide association studies have reported signals of association implicating multiple pathways including host defence, telomere maintenance, signalling and cell-cell adhesion.

Objectives: To improve our understanding of factors that increase IPF susceptibility by identifying previously unreported genetic associations.

Methods and measurements: We conducted genome-wide analyses across three independent studies and meta-analysed these results to generate the largest genome-wide association study of IPF to date (2,668 IPF cases and 8,591 controls). We performed replication in two independent studies (1,456 IPF cases and 11,874 controls) and functional analyses (including statistical fine-mapping, investigations into gene expression and testing for enrichment of IPF susceptibility signals in regulatory regions) to determine putatively causal genes. Polygenic risk scores were used to assess the collective effect of variants not reported as associated with IPF.

Main results: We identified and replicated three new genome-wide significant ($P<5\times10^{-8}$) signals of association with IPF susceptibility (associated with altered gene expression of *KIF15*, *MAD1L1* and *DEPTOR*) and confirmed associations at 11 previously reported loci. Polygenic risk score analyses showed that the combined effect of many thousands of as-yet unreported IPF susceptibility variants contribute to IPF susceptibility.

Conclusions: The observation that decreased *DEPTOR* expression associates with increased susceptibility to IPF, supports recent studies demonstrating the importance of mTOR signalling in lung fibrosis. New signals of association implicating *KIF15* and *MAD1L1* suggest a possible role of mitotic spindle-assembly genes in IPF susceptibility.

Abstract word count: 257

Key words: Genetics, Epidemiology, KIF15, MAD1L1, DEPTOR

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Introduction

Idiopathic pulmonary fibrosis (IPF) is a devastating lung disease characterised by the build-up of scar tissue. It is believed that damage to the alveolar epithelium is followed by an aberrant wound healing response leading to the deposition of dense fibrotic tissue, reducing the lungs' flexibility and inhibiting gas transfer[1]. Treatment options are limited and half of individuals diagnosed with IPF die within 3-5 years[1,2]. Two drugs (pirfenidone and nintedanib) have been approved for the treatment of IPF, but neither offer a cure and only slow disease progression.

IPF is associated with a number of environmental and genetic factors. Identifying regions of the genome contributing to disease risk improves our understanding of the biological processes underlying IPF and helps in the development of new treatments[3]. To date, genome-wide association studies[4-8] (GWAS) have reported 17 common variant (minor allele frequency [MAF]>5%) signals associated with IPF; stressing the importance of host defence, telomere maintenance, cell-cell adhesion and signalling with respect to disease susceptibility. The sentinel (most strongly associated) variant, rs35705950, in one of these signals that maps to the promoter region of the *MUC5B* gene, has a much larger effect on disease susceptibility than other reported risk variants with each copy of the risk allele associated with a five-fold increase in odds of disease[9]. Despite this, the variant rs35705950 has a risk allele frequency of only 35% in cases (compared with 11% in the general population) and so does not explain all IPF risk. Rare variants (MAF<1%) in telomere-related and surfactant genes have also been implicated in familial pulmonary fibrosis and sporadic IPF[10,11].

In this study, we aimed to identify previously unreported genetic associations with IPF to improve our understanding of disease susceptibility and generate new hypotheses about disease pathogenesis. We conducted a large GWAS of IPF susceptibility by utilising all European cases and controls recruited to any previously reported IPF GWAS[5-8] and meta-analysing the results. This was followed by replication in individuals not previously included in IPF GWAS and bioinformatic analysis of gene expression data to identify the genes underlying the identified association signals. As specific IPF associated variants have also been shown to overlap with other related respiratory traits including lung function in the general population, chronic obstructive pulmonary disease (COPD, with genetic effects in opposite directions between COPD and IPF)[12-14] and interstitial lung abnormalities (ILAs, which might be a precursor lesion for IPF)[15], we tested for association of the IPF susceptibility variants with these respiratory phenotypes in independent datasets. Finally, using polygenic risk scores, we tested whether there was a still substantial contribution to IPF risk from genetic variants with as-yet unconfirmed associations with IPF susceptibility.

Some of the results of these studies have been previously reported in the form of an abstract and preprint[16-18].

Methods

Study cohorts

We analysed genome-wide data from three previously described independent IPF case-control collections (named here as the Chicago[5], Colorado[6] and UK[8] studies, please refer to **Appendix** for summaries of these collections). Two more independent case-control collections (named here as the UUS and Genentech studies) were included as replication datasets. The new UUS study recruited cases from the USA, UK and Spain and selected

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

controls from UK Biobank[19] (full details on the recruitment, genotyping and quality control of UUS cases and controls can be found in the **Appendix**). The previously described[20] Genentech study consisted of cases from three IPF clinical trials and controls from four non-IPF clinical trials (**Appendix**). All studies were restricted to unrelated individuals of European ancestry and we applied stringent quality control measures (full details of the quality control measures of each study can be found in the **Appendix** and **Figure E1**). All studies diagnosed cases using American Thoracic Society and European Respiratory Society guidelines[21-23] and had appropriate institutional review board or ethics approval.

Genotype data for the Colorado, Chicago, UK and UUS studies were imputed separately using the Haplotype Reference Consortium (HRC) r1.1 panel[24] (**Appendix**). For individuals in the Genentech study, genotypes were derived from whole-genome sequencing data. Duplicated individuals between studies were removed (**Appendix**).

Identification of IPF susceptibility signals

In each of the Chicago, Colorado and UK studies separately, a genome-wide analysis of IPF susceptibility, using SNPTEST[25] v2.5.2, was conducted adjusting for the first 10 principal components to account for fine-scale population structure. Only bi-allelic autosomal variants that had a minor allele count ≥10, were in Hardy-Weinberg Equilibrium ($P>1\times10^{-6}$), and were well imputed (imputation quality $R^2>0.5$) in at least two studies were included. A genome-wide meta-analysis of the association summary statistics was performed across the Chicago, Colorado and UK studies using R v3.5.1 (discovery stage). Conditional analyses were performed to identify independent association signals in each locus (**Appendix**).

Sentinel variants (defined as the variant in an association signal where no other variants within 1 Mb showed a stronger association) of the novel signals reaching genome-wide

significance in the meta-analysis ($P<5\times10^{-8}$), and nominally significant ($P<0.05$) with consistent direction of effect in each study, were further tested in the replication samples. We considered novel signals to be associated with IPF susceptibility if they reached a Bonferroni-corrected threshold ($P<0.05$ / number of signals followed-up) in a meta-analysis of the UUS and Genentech studies (replication stage, **Appendix**). Previously reported signals with $P<5\times10^{-8}$ in the discovery meta-analysis were deemed as a confirmed association.

Characterisation of signals and functional effects

To further refine our association signals to include only variants with the highest probabilities of being causal, Bayesian fine-mapping was undertaken. This approach takes all variants within the associated locus and, using the GWAS association results, calculates the probability of each variant being the true causal variant (under the assumptions that there is one causal variant and that the causal variant has been measured). The probabilities are then combined across variants to define the smallest set of variants that is 95% likely to contain the causal variant (i.e. the 95% credible set) for each IPF susceptibility signal (**Appendix**).

To identify which genes might be implicated by the IPF susceptibility signals, we identified whether any variants in the credible sets were genic coding variants and defined as deleterious (using VEP[26]). In addition, we tested to see if any of the credible set variants were associated with gene expression using three eQTL resources (the Lung eQTL study [n=1,111][27-29], the NESDA-NTR blood eQTL database [n=4,896][30] and 48 tissues in GTEx[31] [n between 80 and 491], **Appendix**). Where IPF susceptibility variants were found to be associated with expression levels of a gene, we tested whether the same variant was likely to be causal both for differences in gene expression and IPF susceptibility. We only report

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

associations with gene expression where the probability of the same variant driving both the IPF susceptibility signal and gene expression signal exceeded 80% (**Appendix**).

To investigate whether the IPF susceptibility variants that were in non-coding regions of the genome might be in regions with regulatory functions (for example, in regions of open chromatin), we investigated the likely functional impact of those variants using DeepSEA[32]. Taking all of the IPF susceptibility variants together, we tested for overall enrichment in regulatory regions specific to particular cell and tissue types using FORGE[33] and GARFIELD[34]. Finally, we investigated whether the genes that were near to the IPF susceptibility variants were more likely to be differentially expressed between IPF cases and controls in four lung epithelial cell types, using SNPsea[35]. More details are provided in the **Appendix**.

Shared genetic susceptibility with other respiratory traits

As previous studies have reported shared genetic susceptibility for IPF and other lung traits[12,13,15], we investigated whether the new and previously reported IPF susceptibility signals were associated with quantitative lung function measures in a GWAS of 400,102 individuals[36] or with ILA in a GWAS comparing 1,699 individuals with an ILA and 10,247 controls[37]. Lung function measures investigated were, FEV_1 (volume of air an individual can forcibly exhale in the first second), FVC (total volume of air that can be forcibly exhaled), the ratio FEV_1/FVC (used in the diagnosis of COPD) and PEF (the peak expiratory flow). We applied a Bonferroni corrected *P* value threshold to define variants also associated with ILA or lung function.

Polygenic risk scores

The contribution of as-yet unreported variants to IPF susceptibility was assessed using polygenic risk scores. For each individual in the UUS study, the weighted score was calculated as the number of risk alleles, multiplied by the effect size of the variant (as a weighting), summed across all variants included in the score. Effect sizes were taken from the discovery GWAS and independent variants selected using an LD $r^2 \leq 0.1$. As we wanted to explore the contribution from as-yet unreported variants, we excluded variants within 1Mb of each IPF susceptibility locus from the risk score calculation (**Appendix**).

The score was tested to identify whether it was associated with IPF susceptibility, adjusting for 10 principal components to account for fine-scale population structure, using PRSice v1.25[38]. We altered the number of variants included in the risk score calculation using a sliding *P*-threshold (P_T) such that the variant had to have a *P* value$<P_T$ in the genome-wide meta-analysis to be included in the score. This allows us to explore whether variants that do not reach statistical significance in GWAS of current size contribute to disease susceptibility. We used the recommended significance threshold of $P<0.001$ for determining significantly associated risk scores[38].

Data availability statement

Full summary statistics for the genome-wide meta-analysis can be accessed from https://github.com/genomicsITER/PFgenetics.

Results

Following quality control, 541 cases and 542 controls from the Chicago study, 1,515 cases and 4,683 controls from the Colorado study and 612 cases and 3,366 controls from the UK

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

study were available (**Table 1, Figure E1**) to contribute to the discovery stage of the genome-wide susceptibility analysis (**Figure 1**). For the replication stage of the GWAS, after quality control, there were 792 cases and 10,000 controls available in the UUS study and 664 cases and 1,874 controls available in the Genentech study (**Appendix**).

To identify new signals of association, we meta-analysed the genome-wide association results for IPF susceptibility for the Chicago, Colorado and UK discovery studies. This gave a maximum sample size of up to 2,668 cases and 8,591 controls for 10,790,934 well imputed (R^2>0.5) variants with minor allele count ≥10 in each study and which were available in two or more of the studies (**Figure E2**).

Three novel signals (in 3p21.31 [near *KIF15*, **Figure 2i**], 7p22.3 [near *MAD1L1*, **Figure 2ii**] and 8q24.12 [near *DEPTOR*, **Figure 2iii**]) showed a genome-wide significant ($P<5x10^{-8}$) association with IPF susceptibility in the discovery meta-analysis and were also significant after adjusting for multiple testing (*P*<0.01) in the replication stage comprising 1,467 IPF cases and 11,874 controls (**Tables 2 and E1**). Two additional loci were genome-wide significant in the genome-wide discovery analysis but did not reach significance in the replication studies. The sentinel variants of these two signals were a low frequency intronic variant in *RTEL1* (MAF=2.1%, replication *P*=0.012) and a rare intronic variant in *HECTD2* (MAF=0.3%, replication *P*=0.155). Conditional analyses did not identify any additional independent association signals at the new or previously reported IPF susceptibility loci (**Figure E5**).

To identify the likely causal genes for each new signal, we investigated whether any of the variants were also associated with changes in gene expression. The sentinel variant (rs78238620) of the novel signal on chromosome 3 was a low frequency variant (MAF=5%)

in an intron of *KIF15* with the minor allele being associated with increased susceptibility to IPF and decreased expression of *KIF15* in brain tissue and the nearby gene *TMEM42* in thyroid[31] (**Figure E7, Tables E2 and E3i**). The IPF risk allele for the novel chromosome 7 signal (rs12699415, MAF=42%) was associated with decreased expression of *MAD1L1* in heart tissue[31] (**Figure E8, Tables E2 and E3ii**). For the signal on chromosome 8, the sentinel variant (rs28513081) was located in an intron of *DEPTOR* and the IPF risk allele was associated with decreased expression of *DEPTOR* (in colon, lung and skin[27-29,31]) and RP11-760H22.2 (in colon and lung[31]). The risk allele was also associated with *increased* expression of *DEPTOR* (in whole blood[30]), *TAF2* (in colon[31]), RP11-760H22.2 (in adipose[31]) and KB-1471A8.1 (in adipose and skin[31], **Figure E9, Tables E2 and E3iii**). There were no variants predicted to be highly deleterious within the fine-mapped signals for any of the loci.

We confirmed genome-wide significant associations with IPF susceptibility for 11 of the 17 previously reported signals (in or near *TERC, TERT, DSP,* 7q22.1*, MUC5B, ATP11A, IVD, AKAP13, KANSL1, FAM13A* and *DPP9*; **Table E1, Figure E4**). The signal at *FAM13A,* whilst genome-wide significant in the discovery meta-analysis, was not significant in the Chicago study. This was the only signal reaching genome-wide significance in the discovery genome-wide meta-analysis that did not reach at least nominal significance in each study in the discovery analysis. Three further previously reported signals at 11p15.5 (near *MUC5B*) were no longer genome-wide significant after conditioning on the *MUC5B* promoter variant (**Table E1**), consistent with previous reports[6,39].

Of the 14 IPF susceptibility signals (i.e. the 11 previously reported signals we confirmed and three novel signals), the only variant predicted to have a potential functional effect on gene regulation through disruption of chromatin structure or transcription factor binding motifs

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

(using DeepSEA) was rs2013701 (in an intron of *FAM13A*), which was associated with a change in DNase I hypersensitivity in 18 cell types and FOXA1 in the T-47D cell line (a breast cancer cell line derived from a pleural effusion, **Table E4**). The 14 IPF susceptibility signals were found to be enriched in DNase I hypersensitivity site regions in multiple tissues including foetal lung tissue (**Figure E10 and E11**). No enrichment in differential expression in airway epithelial cells between IPF cases and healthy controls was observed for the 14 IPF susceptibility signals when using SNPsea (**Table E5**).

Previous studies have reported an overlap of genetic association loci between lung function and IPF[40]. We undertook a look-up of the 14 IPF susceptibility loci in the largest GWAS of lung function in the general population published to date[36]. The sentinel variants of 12 of the 14 IPF susceptibility loci were at least nominally associated (P<0.05) with one or more lung function trait in general population studies (**Table E6**). After adjustments for multiple testing (P<5.2×10^{-4}), the previously reported variants at *FAM13A*, *DSP* and *IVD* were associated with decreased FVC and variants at *FAM13A*, *DSP*, 7q22.1 (*ZKSCAN1*) and *ATP11A* were associated with increased FEV_1/FVC. Similarly, for the three novel susceptibility variants, all showed at least a nominal association with decreased FVC and increased FEV_1/FVC. We observed a nominally significant association of the *MUC5B* IPF risk allele with decreased FVC and increased FEV_1/FVC. The IPF risk alleles at *MAPT* were significantly associated with both increased FEV_1 and FVC. To determine how the variants identified for IPF susceptibility are related to differences in lung function between cases and controls, we investigated whether variants known to be associated with lung function show an association in our IPF GWAS. Of the 279 variants reported[36] as associated with lung function (**Table E7**), eight showed an association with lung function after corrections for multiple

testing (located in or near *MCL1*, *DSP*, *ZKSCAN1*, *OBFC1*, *IVD*, *MAPT* and two signals in *FAM13A*).

As interstitial lung abnormalities may be a precursor to IPF in a subset of patients, and there have been previous reports of shared genetic aetiology between IPF and ILA[37,41,42], we investigated whether our three new signals, and the 11 previously reported signals, were associated with ILA in the largest ILA GWAS reported to date[37]. Eight of the IPF susceptibility loci were at least nominally significantly associated with either ILA or subpleural ILA with consistent direction of effects (i.e. the allele associated with increased IPF risk was also associated with increased ILA risk). The new *KIF15*, *MAD1L1* and *DEPTOR* signals were not associated with ILA (although the rare risk allele at *HECTD2* that did not replicate in our study showed some association with an increased risk of subpleural ILA [P=0.003] with a large effect size similar to that observed in the IPF discovery meta-analysis).

To quantify the impact of as-yet unreported variants on IPF susceptibility, polygenic risk scores were calculated excluding the 14 IPF susceptibility variants (as well as all variants within 1Mb). The polygenic risk score was significantly associated with increased IPF susceptibility despite exclusion of the known genetic association signals (including *MUC5B*). As the *P*-threshold (P_T) for inclusion of variants in the score was increased, the risk score became more significant reaching a plateau at around P_T=0.2 with risk score $P<3.08\times10^{-23}$ and explaining around 2% of the phenotypic variation (**Figure E12**), suggesting that there is a modest but statistically significant contribution of additional as-yet undetected variants to IPF susceptibility. Further increasing P_T beyond 0.2 did not improve the predictive accuracy of the risk score.

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Discussion

We undertook the largest GWAS of IPF susceptibility to date and identified three novel signals of association that implicated genes not previously known to be important in IPF.

The strongest evidence for the new signal on chromosome 8 implicates *DEPTOR*, which encodes the DEP Domain containing MTOR interacting protein. *DEPTOR* inhibits mTOR kinase activity as part of both the mTORC1 and mTORC2 protein complexes. The IPF risk allele at this locus was associated with decreased gene expression of *DEPTOR* in lung tissue (**Table E2**). TGFβ-induced DEPTOR suppression can stimulate collagen synthesis[43] and the importance of mTORC1 signalling via 4E-BP1 for TGFβ induced collagen synthesis has recently been demonstrated in fibrogenesis[44]. *MAD1L1*, implicated by a new signal on chromosome 7 and eQTL analyses of non-lung tissue, is a mitotic checkpoint gene, mutations in which have been associated with multiple cancers including lung cancer[45,46]. Studies have shown that *MAD1*, a homolog of *MAD1L1*, can inhibit *TERT* activity (or possibly enforce expression of *TERT* when the promoter E-box is mutated)[46,47]. This could suggest that *MAD1L1* may increase IPF susceptibility through reduced telomerase activity. Another spindle-assembly related gene[48], *KIF15,* was implicated by the new signal on chromosome 3 (along with *TMEM42*).

The genome-wide study also identified two signals that were not replicated after multiple testing adjustments. *RTEL1*, a gene involved in telomere elongation regulation has not previously been identified in an IPF GWAS, however the collective effect of rare variants in *RTEL1* have been reported as associated with IPF susceptibility[52-55]. The ubiquitin E3 ligase encoded by *HECTD2* has been shown to have a pro-inflammatory role in the lung and other *HECTD2* variants may be protective against acute respiratory distress syndrome[56]. However,

the lack of replication for these signals in our data suggests that further exploration of their relationship to interstitial lung diseases is warranted.

By combining the largest available GWAS datasets for IPF, we were able to confirm 11 of 17 previously reported signals. Conditional analysis at the 11p15.5 region indicated that previously reported signals at *MUC2* and *TOLLIP* were not independent of the association with the *MUC5B* promoter variant. Previously reported signals at *EHMT2*, *OBFC1* and *MDGA2* were only found to be associated in one of the discovery studies, and showed no evidence of an association with IPF susceptibility in the other two discovery studies. Only the 11 signals that we confirmed in our data were included in subsequent analyses.

The IPF susceptibility signals at *DSP, FAM13A*, 7q22.1 (*ZKSCAN1*) and 17q21.31 (*MAPT*) have also been reported as associated with COPD, although with opposite effects (i.e. the allele associated with increased risk of IPF being associated with decreased risk of COPD). Spirometric diagnosis of COPD was based on a reduced FEV_1/FVC ratio. In an independent dataset of 400,102 individuals, eight of the IPF signals were associated with decreased FVC and with a comparatively weaker effect on FEV_1. This is consistent with the lung function abnormalities associated with IPF, as well as the decreased risk of COPD. Of note, only around 3% of previously reported lung function signals[36] also showed association with IPF susceptibility in our study. This suggests that whilst some IPF susceptibility variants might represent genes and pathways that are important in general lung health, others are likely to represent more disease-specific processes.

Using polygenic risk scores, we demonstrated that, despite the relatively large proportion of disease susceptibility explained by the known genetic signals of association reported here,

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

IPF is highly polygenic with potentially hundreds (or thousands) of as-yet unidentified variants associated with disease susceptibility.

A strength of our study was the large sample size compared with previous GWAS and the availability of an independent replication data set. A limitation of our study was that the controls used were generally younger in all studies included and there were differences in sex and smoking distributions in some of the studies. As age, sex and smoking status were not available for all individuals in four of our datasets, we were unable to adjust for these variables without substantially reducing our sample size. However, cases and controls in the UUS and UK datasets were matched for age, sex and smoking. The three novel signals replicated in all of the discovery and replication datasets providing reassurance that the signals we report are robust despite differences between the data sets. As we had limited information beyond IPF diagnosis status for a large proportion of the individuals included in the studies, we cannot rule out some association with other age-related conditions that are comorbid with IPF. However, other age-related conditions were not excluded from either the cases or controls. For the signals near *KIF15* and *MAD1L1*, there was substantial evidence for an association with gene expression in non-lung tissues but not in either of the two (non-fibrotic) lung tissue eQTL datasets. This could reflect cell type-specific effects that are missed when studying whole tissue or effects that are disease dependent. Finally, our study was not designed to identify rare functional variant associations. As both common and rare variants are known to be important in IPF susceptibility[39], this is a limitation of our study.

In summary, we report new biological insights into IPF susceptibility and demonstrate that further studies to identify the genetic determinants of IPF susceptibility are needed. Our

new signals of association with IPF susceptibility provide increased support for the importance of mTOR signalling in pulmonary fibrosis as well as the possible implication of mitotic spindle-assembly genes.

Acknowledgements

This research has been conducted using the UK Biobank Resource under application 8389. This research used the ALICE and SPECTRE High Performance Computing Facilities at the University of Leicester. Genome-wide summary statistics are available on request via the corresponding author.

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

References

1. Lederer DJ, Martinez FJ. Idiopathic pulmonary fibrosis. *N Engl J Med*. 2018;378(19):1811-1823.

2. Ley B, Collard HR, King Jr TE. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*. 2011;183(4):431-440.

3. Nelson MR, Tipney H, Painter JL, et al. The support of human genetic evidence for approved drug indications. *Nat Genet*. 2015;47(8):856-860.

4. Mushiroda T, Wattanapokayakit S, Takahashi A, et al. A genome-wide association study identifies an association of a common variant in TERT with susceptibility to idiopathic pulmonary fibrosis. *J Med Genet*. 2008;45(10):654-656.

5. Noth I, Zhang Y, Ma S, et al. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: A genome-wide association study. *The Lancet Respiratory Medicine*. 2013;1(4):309-317.

6. Fingerlin TE, Murphy E, Zhang W, et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet*. 2013;45(6):613-620.

7. Fingerlin TE, Zhang W, Yang IV, et al. Genome-wide imputation study identifies novel HLA locus for pulmonary fibrosis and potential role for auto-immunity in fibrotic idiopathic interstitial pneumonia. *BMC genetics*. 2016;17(1):1.

8. Allen RJ, Porte J, Braybrooke R, et al. Genetic variants associated with susceptibility to idiopathic pulmonary fibrosis in people of european ancestry: A genome-wide association study. *The Lancet Respiratory Medicine*. 2017.

9. Zhu QQ, Zhang XL, Zhang SM, et al. Association between the MUC5B promoter polymorphism rs35705950 and idiopathic pulmonary fibrosis: A meta-analysis and trial sequential analysis in caucasian and asian populations. *Medicine (Baltimore)*. 2015;94(43):e1901.

10. Coghlan MA, Shifren A, Huang HJ, et al. Sequencing of idiopathic pulmonary fibrosis-related genes reveals independent single gene associations. *BMJ open respiratory research*. 2014;1(1):e000057.

11. Petrovski S, Todd JL, Durheim MT, et al. An exome sequencing study to assess the role of rare genetic variation in pulmonary fibrosis. *American journal of respiratory and critical care medicine*. 2017;196(1):82-93.

12. Hobbs BD, De Jong K, Lamontagne M, et al. Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nat Genet*. 2017;49(3):426.

13. Sakornsakolpat P, Prokopenko D, Lamontagne M, et al. Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. *Nat Genet*. 2019;51:494-505.

14. Chilosi M, Poletti V, Rossi A. The pathogenesis of COPD and IPF: Distinct horns of the same devil. *Respir Res*. 2012;13(3).

15. Putman RK, Gudmundsson G, Araki T, et al. The MUC5B promoter polymorphism is associated with specific interstitial lung abnormality subtypes. *Eur Respir J*. 2017;50(3):10.1183/13993003.00537-2017. Print 2017 Sep.

16. Allen RJ, Oldham JM, Fingerlin TE, et al. Polygenicity of idiopathic pulmonary fibrosis [abstract]. *Genetic Epidemiology*. 2018;42:684.

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

17. Allen RJ, Guillen-Guio B, Oldham JM, et al. Genome-wide meta-analysis of susceptibility to idiopathic pulmonary fibrosis [abstract]. *American Society of Human Genetics*. 2018.

18. Allen RJ, Guillen-Guio B, Oldham JM, et al. Genome-wide association study of susceptibility to idiopathic pulmonary fibrosis. *bioRxiv*. 2019:636761.

19. Sudlow C, Gallacher J, Allen N, et al. UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;12(3):e1001779.

20. Dressen A, Abbas AR, Cabanski C, et al. Analysis of protein-altering variants in telomerase genes and their association with MUC5B common variant status in patients with idiopathic pulmonary fibrosis: A candidate gene sequencing study. *The Lancet Respiratory Medicine*. 2018.

21. Travis WD, King TE, Bateman ED, et al. American thoracic society/european respiratory society international multidisciplinary consensus classification of the idiopathic interstitial pneumonias. *American journal of respiratory and critical care medicine*. 2002;165(2):277-304.

22. Raghu G, Collard HR, Egan JJ, et al. An official ATS/ERS/JRS/ALAT statement: Idiopathic pulmonary fibrosis: Evidence-based guidelines for diagnosis and management. *American Journal of Respiratory and Critical Care Medicine*. 2011;183(6):788-824.

23. Raghu G, Remy-Jardin M, Myers JL, et al. Diagnosis of idiopathic pulmonary fibrosis. an official ATS/ERS/JRS/ALAT clinical practice guideline. *American journal of respiratory and critical care medicine*. 2018;198(5):e44-e68.

24. McCarthy S, Das S, Kretzschmar W, Durbin R, Abecasis G, Marchini J. A reference panel of 64,976 haplotypes for genotype imputation. *bioRxiv*. 2016:035170.

25. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*. 2007;39(7):906-913.

26. McLaren W, Gil L, Hunt SE, et al. The ensembl variant effect predictor. *Genome Biol*. 2016;17(1):122.

27. Hao K, Bossé Y, Nickle DC, et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS Genet*. 2012;8(11):e1003029.

28. Lamontagne M, Couture C, Postma DS, et al. Refining susceptibility loci of chronic obstructive pulmonary disease with lung eqtls. *PLoS One*. 2013;8(7):e70220.

29. Obeidat M, Miller S, Probert K, et al. GSTCD and INTS12 regulation and expression in the human lung. *PLoS One*. 2013;8(9):e74630.

30. Jansen R, Hottenga J, Nivard MG, et al. Conditional eQTL analysis reveals allelic heterogeneity of gene expression. *Hum Mol Genet*. 2017;26(8):1444-1451.

31. GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature*. 2017;550(7675):204.

32. Zhou J, Troyanskaya OG. Predicting effects of noncoding variants with deep learning–based sequence model. *Nature methods*. 2015;12(10):931.

33. Dunham I, Kulesha E, Iotchkova V, Morganella S, Birney E. FORGE: A tool to discover cell specific enrichments of GWAS associated SNPs in regulatory regions. *F1000Research*. 2015;4.

34. Iotchkova V, Ritchie GR, Geihs M, et al. GARFIELD-GWAS analysis of regulatory or functional information enrichment with LD correction. *bioRxiv*. 2016:085738.

35. Slowikowski K, Hu X, Raychaudhuri S. SNPsea: An algorithm to identify cell types, tissues and pathways affected by risk loci. *Bioinformatics*. 2014;30(17):2496-2497.

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

36. Shrine N, Guyatt AL, Erzurumluoglu AM, et al. New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nat Genet*. 2019:1.

37. Hobbs BD, Putman RK, Araki T, et al. Overlap of genetic risk between interstitial lung abnormalities and idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*. 2019((in press)).

38. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic risk score software. *Bioinformatics*. 2015;31(9):1466-1468.

39. Moore C, Blumhagen R, Yang I, et al. Resequencing study confirms host defence and cell senescence gene variants contribute to the risk of idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*. 2019.

40. Hobbs BD, De Jong K, Lamontagne M, et al. Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nat Genet*. 2017;49(3):426.

41. Hunninghake GM, Hatabu H, Okajima Y, et al. MUC5B promoter polymorphism and interstitial lung abnormalities. *N Engl J Med*. 2013;368(23):2192-2200.

42. Putman RK, Hunninghake GM, Dieffenbach PB, et al. Interstitial lung abnormalities are associated with acute respiratory distress syndrome. *American journal of respiratory and critical care medicine*. 2017;195(1):138-141.

43. Das F, Bera A, Ghosh-Choudhury N, Abboud HE, Kasinath BS, Choudhury GG. TGFβ-induced deptor suppression recruits mTORC1 and not mTORC2 to enhance collagen I (α2) gene expression. *PloS one*. 2014;9(10):e109608.

44. Woodcock HV, Eley JD, Guillotin D, et al. The mTORC1/4E-BP1 axis represents a critical signaling node during fibrogenesis. *Nature communications*. 2019;10(1):6.

45. Thompson MJ, Rubbi L, Dawson DW, Donahue TR, Pellegrini M. Pancreatic cancer patient survival correlates with DNA methylation of pancreas development genes. *PLoS One*. 2015;10(6):e0128814.

46. Coe BP, Lee EH, Chi B, et al. Gain of a region on 7p22. 3, containing MAD1L1, is the most frequent event in small-cell lung cancer cell lines. *Genes, Chromosomes and Cancer*. 2006;45(1):11-19.

47. Kang JU, Koo SH, Kwon KC, Park JW, Kim JM. Gain at chromosomal region 5p15. 33, containing TERT, is the most frequent genetic event in early stages of non-small cell lung cancer. *Cancer Genet Cytogenet*. 2008;182(1):1-11.

48. Tanenbaum ME, Macůrek L, Janssen A, Geers EF, Alvarez-Fernández M, Medema RH. Kif15 cooperates with eg5 to promote bipolar spindle assembly. *Current biology*. 2009;19(20):1703-1711.

49. Alder JK, Chen JJ, Lancaster L, et al. Short telomeres are a risk factor for idiopathic pulmonary fibrosis. *Proc Natl Acad Sci U S A*. 2008;105(35):13051-13056.

50. Stuart BD, Lee JS, Kozlitina J, et al. Effect of telomere length on survival in patients with idiopathic pulmonary fibrosis: An observational cohort study with independent validation. *The lancet Respiratory medicine*. 2014;2(7):557-565.

51. McDonough JE, Martens DS, Tanabe N, et al. A role for telomere length and chromosomal damage in idiopathic pulmonary fibrosis. *Respiratory research*. 2018;19(1):132.

52. Stuart BD, Choi J, Zaidi S, et al. Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. *Nat Genet*. 2015;47(5):512-517.

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

53. Kropski JA, Loyd JE. Telomeres revisited: RTEL1 variants in pulmonary fibrosis. *Eur Respir J*. 2015;46(2):312-314.

54. Kannengiesser C, Borie R, Menard C, et al. Heterozygous RTEL1 mutations are associated with familial pulmonary fibrosis. *Eur Respir J*. 2015;46(2):474-485.

55. Deng Y, Li Z, Liu J, et al. Targeted resequencing reveals genetic risks in patients with sporadic idiopathic pulmonary fibrosis. *Hum Mutat*. 2018;39(9):1238-1245.

56. Coon TA, McKelvey AC, Lear T, et al. The proinflammatory role of HECTD2 in innate immunity and experimental lung injury. *Science translational medicine*. 2015;7(295):295ra109-295ra109.

57. Dudbridge F. Power and predictive accuracy of polygenic risk scores. *PLoS genetics*. 2013;9(3):e1003348.

58. Calafato MS, Thygesen JH, Ranlund S, et al. Use of schizophrenia and bipolar disorder polygenic risk scores to identify psychotic disorders. *The British Journal of Psychiatry*. 2018;213(3):535-541.

59. Mullins N, Power R, Fisher H, et al. Polygenic interactions with environmental adversity in the aetiology of major depressive disorder. *Psychol Med*. 2016;46(4):759-770.

60. Wain LV, Shrine N, Miller S, et al. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): A genetic association study in UK biobank. *The Lancet Respiratory Medicine*. 2015;3(10):769-781.

Table 1: Demographics of study cohorts

	Chicago		Colorado		UK		UUS		Genentech	
	Cases	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**	**Cases**	**Controls**
n	541	542	1,515	4,683	612	3,366	792	10,000	664	1,874
Genotyping array /sequencing	Affymetrix 6.0 SNP array		Illumina Human 660W Quad BeadChip		Affymetrix UK BiLEVE array	Affymetrix UK BiLEVE and UK Biobank arrays	Affymetrix UK Biobank and Spain Biobank arrays	Affymetrix UK BiLEVE and UK Biobank arrays	HiSeq X Ten platform (Illumina)	
Imputation panel	HRC		HRC		HRC		HRC		-	
Age (mean)	68	63 [a]	66	-	70 [b]	65	69	58	68	-
Sex (% males)	71% [c]	47% [d]	68%	49%	70.8%	70.0%	75.2%	72.1%	73.5%	27.1%
% ever smokers	72%	42%	-	-	72.9% [e]	70.0%	68.7% [f]	68.0%	67.3%	18.1% [g]

[a] Age only available for 103 Chicago controls
[b] Age available for 602 UK cases
[c] Sex only available for 500 Chicago cases
[d] Sex only available for 510 Chicago controls
[e] Smoking status only recorded for 236 UK cases
[f] Smoking status only recorded for 753 IPF cases in UUS
[g] Smoking status only recorded for 481 of the Genentech controls

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Table 2 - Discovery and replication association analysis results for the five signals reaching significance in the discovery GWAS that have not previously reported as associated with IPF

The minor allele is the effect allele and the minor allele frequency (MAF) is taken from across the studies used in the discovery meta-analysis.

Chr	Pos	rsid	Locus	Major allele	Minor allele	MAF	Discovery meta-analysis		Replication meta-analysis		Meta-analysis of discovery and replication	
							OR [95% CI]	*P*	OR [95% CI]	*P*	OR [95% CI]	*P*
3	44902386	rs78238620	*KIF15*	T	A	5.3%	1.58 [1.37, 1.83]	5.12×10^{-10}	1.48 [1.24, 1.77]	1.43×10^{-5}	1.54 [1.38, 1.73]	4.05×10^{-14}
7	1909479	rs12699415	*MAD1L1*	G	A	42.0%	1.28 [1.19, 1.37]	7.15×10^{-13}	1.29 [1.18, 1.41]	2.27×10^{-8}	1.28 [1.21, 1.35]	9.38×10^{-20}
8	120934126	rs28513081	*DEPTOR*	A	G	42.8%	0.82 [0.76, 0.87]	1.20×10^{-9}	0.87 [0.80, 0.95]	0.002	0.83 [0.79, 0.88]	1.93×10^{-11}
10	93271016	rs537322302	*HECTD2*	C	G	0.3%	7.82 [3.77, 16.2]	3.43×10^{-8}	1.75 [0.81, 3.78]	0.155	3.85 [2.27, 6.54]	6.25×10^{-7}
20	62324391	rs41308092	*RTEL1*	G	A	2.1%	2.12 [1.67, 2.69]	7.65×10^{-10}	1.45 [1.08, 1.94]	0.012	1.82 [1.51, 2.19]	2.24×10^{-10}

Table 3 – Gene expression and spirometric results for the three novel IPF susceptibility loci

Annotation of the variant was taken from VEP. A list of all variants included in the credible sets with their annotations and eQTL results can be found in **Table E3**. For colocalisation, only genes where there was a greater than 80% probability of colocalisation between the IPF risk signal and gene expression of that gene are reported in this table. In the colocalisation column, ↑ denotes that the allele that increases IPF risk was associated with increased expression of the gene, ↓ denotes that the IPF risk allele was associated with decreased expression of the gene and ↕ denotes that the IPF risk allele was associated with increased expression in some tissues and decreased expression in others. Full results from the eQTL and colocalisation analyses can be found in **Table E2**. The spirometric results for the three novel IPF risk loci are taken from Shrine et al using the allele associated with increased IPF risk as the effect allele with β being the change in Z-score units. Results for all IPF risk variants can be found in **Table E6**.

Chr	rsid of sentinel variant	Annotation	eQTL		FEV_1		FVC		FEV_1 / FVC	
			Lung tissue	Non-lung tissue	β [95% CI]	*P*	β [95% CI]	*P*	β [95% CI]	*P*
3	rs78238620	Intron (*KIF15*)	-	↓ *KIF15* ↓ *TMEM42*	−0.011 [−0.022, 0.000]	0.069	−0.022 [-0.033, 0.011]	2.92×10^{-4}	0.017 [0.006, 0.028]	0.005
7	rs12699415	Intron (*MAD1L1*)	-	↓ *MAD1L1*	−0.007 [−0.012, −0.002]	0.011	−0.011 [−0.016, −0.007]	1.41×10^{-5}	0.008 [0.003, 0.012]	0.005
8	rs28513081	Intron (*DEPTOR*)	↓ *DEPTOR* ↓ RP11-760H22.2	↕ *DEPTOR* ↕ RP11-760H22.2 ↑ KB-1471A8.1 ↑ *TAF2*	0.001 [−0.004, 0.006]	0.822	−0.005 [−0.010, −0.001]	0.045	0.011 [0.006, 0.016]	4.22×10^{-5}

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Figure 1 - Manhattan plot of discovery analysis results

X axis shows chromosomal position and the y axis shows the −log(*P* value) for each variant in the discovery genome-wide analysis. The red line shows genome-wide significance ($P<5\times10^{-8}$) and variants in green met the criteria for further study in the replication analysis (i.e. reached genome-wide significance in the discovery meta-analysis and had *P*<0.05 and consistent direction of effects in each study). Genes labelled in grey are previously reported signals that reach significance in the discovery genome-wide meta-analysis. Genes labelled in black are the novel signals identified in the discovery analysis that reach genome-wide significance when meta-analysing discovery and replication samples. The signals which did not replicate are shown by red labels. For ease of visualisation the y axis has been truncated at 25.

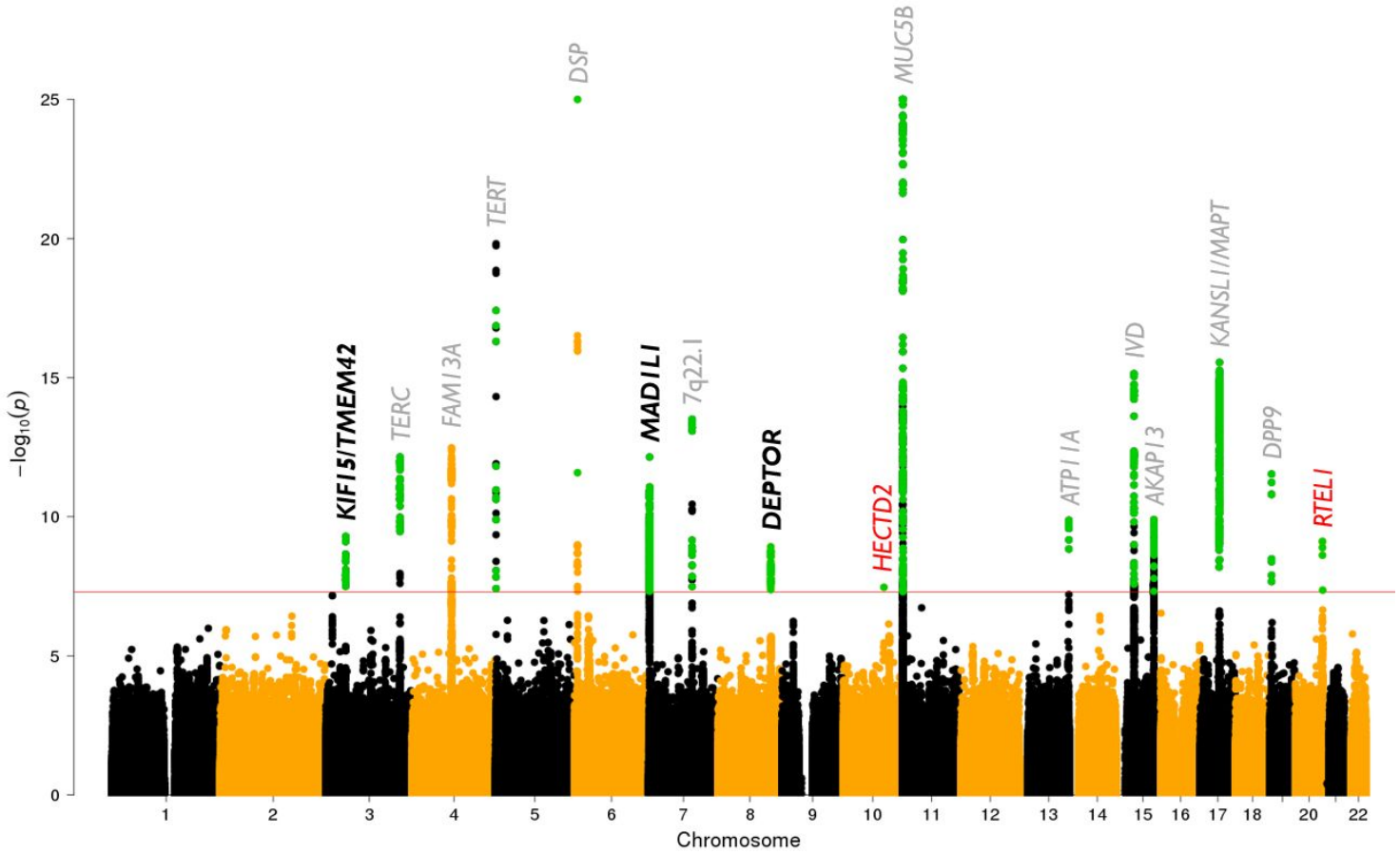
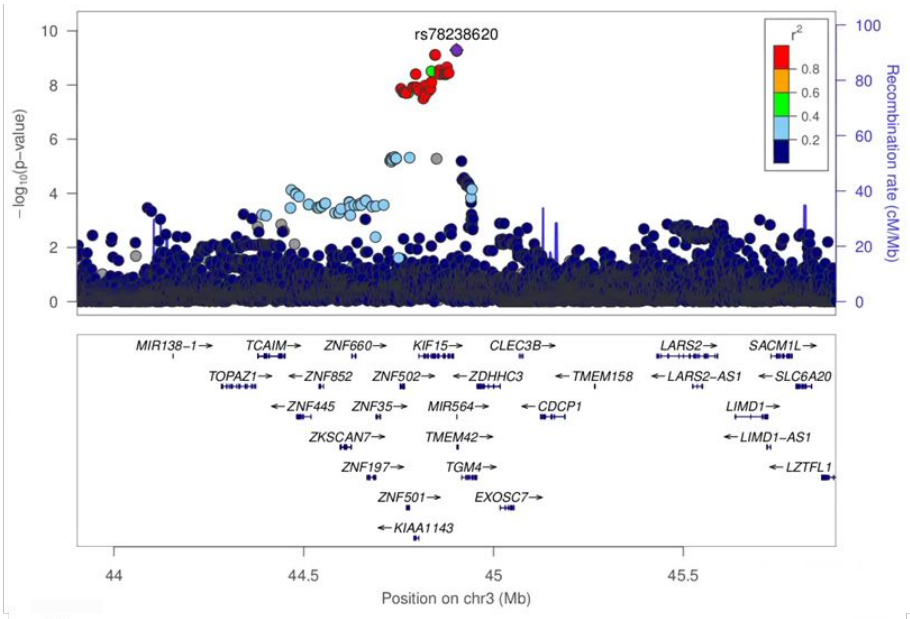


Figure 2 - Region plots of three novel IPF susceptibility loci from discovery genome-wide meta-analysis

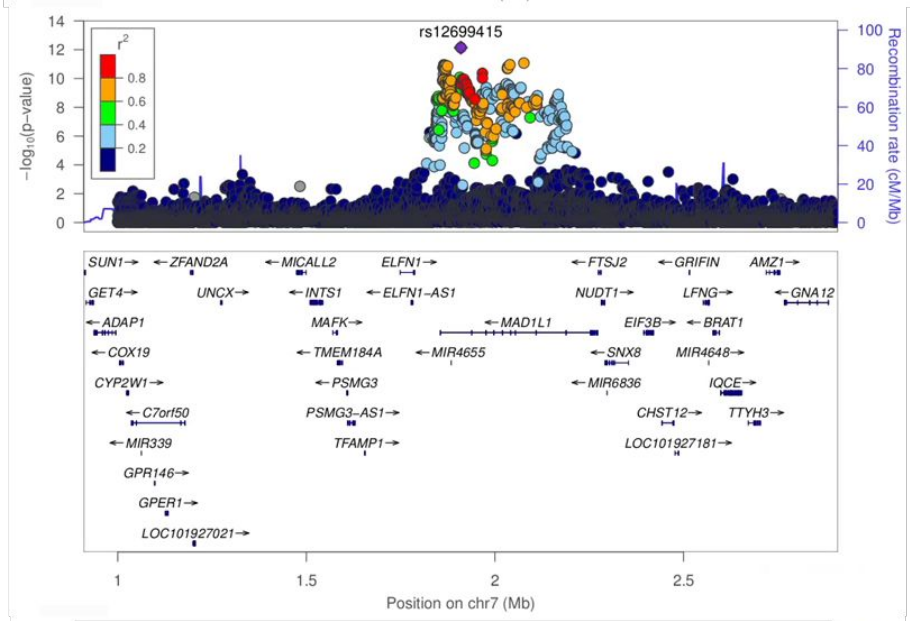
Each point represents a variant with chromosomal position on the x axis and the −log(*P* value) on the y axis. Variants are coloured in by LD with the sentinel variant. Blue lines show the recombination rate and gene locations are shown at the bottom of the plot. Region plots are shown for the three replicated novel IPF susceptibility loci, i.e. i) the susceptibility signal on chromosome 3 near *KIF15*, ii) the susceptibility signal on chromosome 7 near *MAD1L1* and iii) the susceptibility signal on chromosome 8 near *DEPTOR*.

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

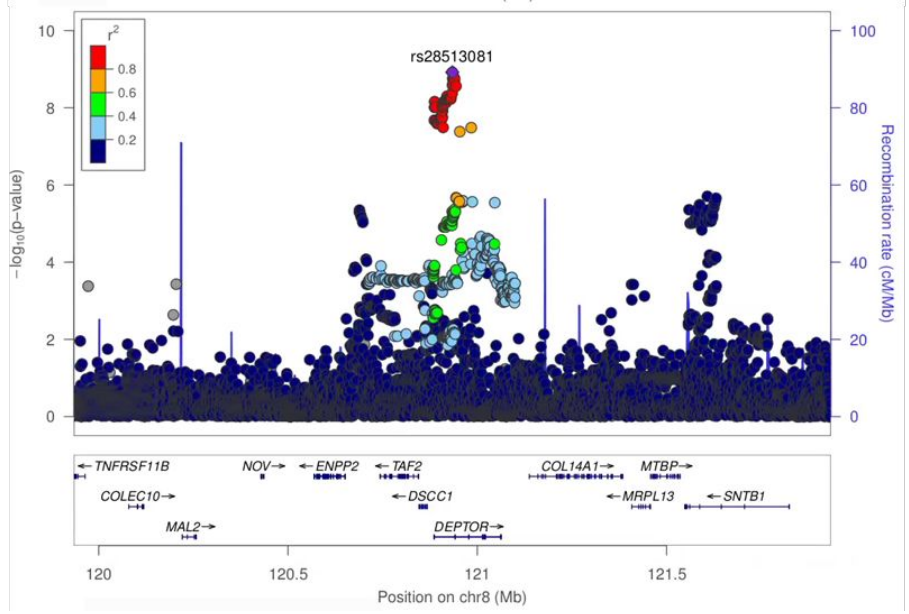
i)



ii)



iii)



Appendix

Online Data Supplement

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Supplementary Methods

Overview of study

In this study we analysed data from five different idiopathic pulmonary fibrosis (IPF) studies, which in this manuscript we refer to as the Chicago, Colorado, UK, UUS and Genentech studies. Three of these studies (Chicago, Colorado and UK) have been used for previous genome-wide association studies (GWAS)[1-4]. The Genentech study consisted of IPF cases and controls from clinical trials and have been previously described[5]. The UUS study consisted of newly genotyped cases and is described more fully in this supplement.

We reimputed the cases and controls in the Chicago, Colorado and UK studies and reran genome-wide analyses in each of these three studies separately and meta-analysed the results to perform the largest most powerful IPF GWAS to date. Novel associations identified in this meta-analysis were then tested for an association in the Genentech and UUS studies. Functional follow-up analyses were used to determine putatively causal genes.

Summary of previously reported studies

The Chicago study[1] consisted of 541 IPF cases and 542 controls. Cases were selected from the University of Chicago, University of Pittsburgh and COMET study and the controls were selected from the database of genotypes and phenotypes (dbGaP) and healthy individuals recruited from the University of Pittsburgh. All individuals were unrelated, of European-American ancestry and had a genotyping call rate > 97%. Subjects with sex mismatches were removed and controls were selected so they were genetically matched to a case based on the first 4 principal components. All individuals were genotyped using the Genome-Wide Human SNP Array 6.0 (Affymetrix).

The Colorado Study[2,3] consisted of 1,616 fibrotic IIP (idiopathic interstitial pneumonia) cases (from the National Jewish Health IIP population, InterMune IPF trials, UCSF, Vanderbilt University IIP

population and the National Heart, Lung and Blood Institute Lung Tissue Research Consortium) and 4,683 controls (generated at Centre d'Etude du Polymorphisme Humain and approved for use as controls in other studies). Controls were selected such that they were genetically similar to the cases based on IBS (identical by state) estimates. Of the cases, 101 were removed from this study as they had also been included in the Chicago study (see **Duplicated individuals between studies** section). All individuals were self-reported as non-Hispanic white and were removed if they had a genotyping call rate < 98%, were sex mismatches, had genome-wide heterozygosity more than four standard deviations away from the mean or were genetic outliers based on IBS estimates. All individuals were genotyped using the Human 660W Quad BeadChip array (Illumina Inc.).

The UK study[4] consisted of 612 IPF cases and 3,366 controls selected from UK Biobank such that they had no history of any interstitial lung disease (defined by hospital episode statistics and cause of death) and followed a similar age, sex and smoking distribution to the cases. Individuals were removed if they had high missingness (call rate < 95%), were heterozygote outliers, were ancestry outliers based on principal components or were sex mismatches. All individuals were of European ancestry and were unrelated. Genotyping of cases was performed using the UK BiLEVE array (Affymetrix). For the controls, 1,231 were genotyped using the UK BiLEVE array and the remaining 2,135 were genotyped using the similar UK Biobank array (Affymetrix).

The Genentech study[5] consisted of 664 unrelated European IPF cases taken from the ASCEND, CAPACITY and RIFF clinical trials and 1,874 unrelated European non-IPF controls taken from the EXCELS, SUMMACTA, LITHE and OPTION clinical trials. The original Genentech cohort also included IPF cases from the Vanderbilt, UCSF and INSPIRE cohorts, however as these had been included in other studies included in the discovery analysis, these individuals were excluded. Individuals were sequenced using the HiSeq X Ten platform (Illumina Inc.) to an average read depth of 30X. Individuals were excluded from analyses if they had a call rate < 10%, had excess heterozygosity, were ancestry outliers or were aged less than 40 years old.

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Recruitment and genotyping of cases for UUS (USA, UK and Spain) study

A total of 1,288 individuals were recruited across the USA, UK and Spain and sent for genotyping from 7 study cohorts, namely; ACE (Anticoagulant Effectiveness in IPF, n = 98), PANTHER (Prednisone, Azathioprine, and N-Acetylcysteine: A Study That Evaluates Response in Idiopathic Pulmonary Fibrosis, n = 166), UCD (University of California Davis, n = 54), Chicago (n = 314), UCSF (University of California San Francisco, n = 53), PROFILE (n = 554) and Spain (n = 50). These collections were primarily intended to capture sporadic cases though family history was not recorded for all study cohorts.

Cases in the replication study were genotyped by Affymetrix on the Axiom UK Biobank array, apart from the 50 Spanish IPF cases who were genotyped on the Axiom Spain Biobank array (Affymetrix). The UK Biobank array was designed to optimise imputation quality of common (MAF [minor allele frequency] > 5%) and low-frequency (MAF 1% to 5%) variants in a European population, measure rare functional variation and to include custom content of known genetic associations with a variety of traits (including *MUC5B* promoter and *TERT* variants with known associations with IPF).

Quality control was performed on the individuals on the two arrays separately before being merged for the selection of controls and for imputation.

Quality control for UUS study

Quality control for individuals genotyped on UK Biobank array

For the individuals genotyped on the UK Biobank array, the following quality control measures were applied.

1. **Affymetrix quality control:** Individuals were removed if they had scanning issues, failed dish QC and or had a sample call rate < 97% in step 1 genotype calling. The genotype calling and quality control was originally performed by Affymetrix and was repeated using Axiom

Analysis Suite and APT (Analysis Power Tools). All three calling methods gave the same results.

2. **Individual call rate:** Individuals were excluded for having a final individual call rate < 95%.
3. **Sex mismatches:** Genetic sex was inferred using PLINK v1.9. Individuals who had a recorded sex different to that inferred from their genetic sample and were not included in further analyses.
4. **High heterozygosity:** Individuals were excluded if they had a high heterozygosity rate (defined as more than 5 standard deviations above the mean after adjusting for ancestry). Heterozygosity rates were calculated using autosomal variants in Hardy-Weinberg equilibrium with MAF > 1% and variant call rate > 95%.
5. **Non-IPF cases:** Individuals found to not be IPF cases were removed.
6. **Duplicates:** Duplicates were identified using KING on all samples, including those individuals already excluded for failing other quality control measures. The duplicate analysis was performed on autosomal variants in Hardy-Weinberg equilibrium ($P > 10^{-6}$), had call rate > 95%, MAF > 1% and not found in regions of high linkage disequilibrium (LD, namely positions 44Mb to 51.5Mb on chromosome 5, 25Mb to 33.5Mb on chromosome 6 [HLA region], 8Mb to 12Mb on chromosome 8 and 45Mb to 57Mb on chromosome 11). If the phenotype data suggested the same person had been recruited twice then the sample with the highest call rate was kept. In instances where the phenotype data suggested there had been a potential genetic sample mix-up, both pairs were removed. Duplicates were also identified between studies. More details on this analysis can be found in the "Duplicated individuals between studies" section in the supplementary methods.
7. **Ancestry:** Ancestry was inferred from the genetic data using principal components analysis. Principal components were calculated using PLINK v1.9 on autosomal variants with MAF > 1%, in Hardy-Weinberg equilibrium, were variants included in HapMap, had genotyping call rate > 95% and were not in regions of high LD. Variants were pruned using an

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

r^2 threshold of 0.1. Principal components were calculated for the individuals who passed Affymetrix QC, had call rate > 95%, were not sex mismatches and were not duplicates and for all unrelated samples from the HapMap project (a collection of genotyped individuals from multiple populations). K-means clustering on the first two principal components was used to define ancestry groups. The number of clusters was increased until a cluster was formed which contained all European HapMap samples and no HapMap samples of other ancestries. Using seven clusters was found to form a cluster of European samples. Individuals in the other ancestry clusters were not included in further analyses.

8. **Relatedness:** Relatedness between individuals passing quality control measures was calculated using KING. First-degree relatives were defined as those with kingship coefficient between 0.177 and 0.354 and second-degree relatives as a kingship coefficient between 0.0884 and 0.177. When first and second-degree relatives were identified, the individual with the lower genotyping call rate was removed from further analyses.
9. **Relatedness with the discovery:** Individuals who were first or second-degree relatives with an individual in the discovery analysis were also excluded. Relatedness was estimated using KING.

Of the 1,238 individuals genotyped on the UK Biobank array, 753 unrelated European IPF cases passed quality control.

Quality control of individuals genotyped on Spain Biobank array for UUS study

For the individuals genotyped on the Spanish custom array, the following quality control measures were applied.

1. **Affymetrix QC:** Individuals were removed if they had scanning issues, failed dish QC and or had a sample call rate < 97% in step 1 genotype calling. The genotype calling and quality control was performed using Axiom Analysis Suite.

2. **Individual call rate:** Individuals were excluded for having a final individual call rate < 95%.
3. **Sex mismatches:** Sex was inferred from the genetic data using PLINK v1.9. Individuals were excluded from future analyses if their genetically inferred sex was different to their recorded sex.
4. **High heterozygosity:** Heterozygosity rates were estimated using PLINK v1.9 on autosomal variants with call rate > 95%, in Hardy-Weinberg equilibrium ($P > 10^{-6}$) and MAF > 1%. Heterozygosity rates were adjusted for ancestry and individuals found to have a high genome-wide ancestry-adjusted heterozygosity rate (more than 5 standard deviations above mean) were removed.
5. **Duplicates and relatedness:** Duplicates and relatedness between individuals was calculated using PLINK v1.9 on autosomal variants with call rate > 95%, in Hardy-Weinberg equilibrium ($P > 10^{-6}$), MAF > 1% and not in a region of high LD. Variants were pruned using an r^2 threshold of 0.13. In instances of duplicates, first or second-degree relatives the sample with the lowest genotyping call rate was removed (apart from instances where it appeared genetic sample mix-up had occurred in which case both individuals were removed).
6. **Ancestry:** Ancestry outliers for the IPF cases passing previous quality control measures were inferred from the genetic data through principal components analysis. Principal components were calculated using PLINK v1.9 on autosomal variants with MAF > 1%, in Hardy-Weinberg equilibrium, were variants included in HapMap, had genotyping call rate > 95% and were not in regions of high LD. Variants were pruned using an r^2 threshold of 0.22 (the lowest r^2 value that left more than 100,000 variants). Principal components were calculated alongside HapMap samples. Individuals who had either the first or second principal component greater than two standard deviations away from the mean were deemed to be ancestry outliers and removed from further analyses.

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

7. **Relatedness with the discovery:** Individuals who were first or second-degree relatives with an individual in the discovery analysis were also excluded. Relatedness was estimated using KING.

Of the 50 individuals genotyped on the Spanish custom array, 39 unrelated European IPF cases passed quality control.

Selection and quality control of controls for UUS study

Controls were selected from UK Biobank such that they were European (defined by k-means clustering of first two principal components), not a possible ILD case, were related to another UK Biobank individual, or were a control in the UK study.

ILD cases in UK Biobank were identified using the self-reported questionnaire (field 20002 - Non-cancer illness code, self-reported) and from HES data (i.e. any hospital episode recorded with ICD10 codes J84, J841, J848, J849 or ICD9 codes 516, 5160, 5161, 5162, 5163, 51630, 51631, 51632, 51633, 51634, 51635, 51636, 51637, 5164, 5165, 5166, 51661, 51662, 51663, 51664, 51669, 5168, 5169).

Of the 300,909 individuals passing the above selection criteria, 10,000 were selected as controls such that they followed a similar sex and smoking distribution to that seen in the IPF cases.

Imputation of all studies

Each study was imputed separately to the Haplotype Reference Consortium reference panel using the Michigan Imputation Server. Only variants in Hardy-Weinberg equilibrium ($P > 10^{-6}$), had call rate > 95% and had MAF > 1% were considered.

When more than one genotyping array was used in a study (i.e. the UK study and replication study) only variants that appeared on all arrays used in that study were included in the imputation. For the replication study, the concordance between the imputed genotypes and the directly measured

genotypes not included in the imputation (i.e. due to not being on all the arrays used) was found to be high (concordance = 99.6%).

Duplicated individuals between studies

It is possible for individuals to be recruited to multiple studies. To ensure the studies included in this analysis were completely independent, individuals who had been recruited to multiple studies were identified from the genetic data and removed. This was conducted using PLINK v1.9 and verified using KING.

Variants were included in the PLINK IBD (identical by descent) analysis if they were on an autosome, had call rate > 95%, MAF > 1%, in Hardy-Weinberg equilibrium ($P > 10^{-6}$) and not in a region of high LD. Variants were pruned using an r^2 threshold of 0.3 leaving 120,864 variants to be included in the IBD analysis (as a sensitivity analysis an r^2 threshold of 0.1 was used and the same results were observed). Pairs of genetic samples with PI_HAT > 0.8 were considered as duplicates.

The duplicate analysis was repeated using KING on autosomal variants with call rate > 95%, MAF > 1%, in Hardy-Weinberg equilibrium ($P > 10^{-6}$) and not in an area of high linkage disequilibrium. Duplicates were identified as those with kingship > 0.354. The KING analysis gave the same results as seen in the analysis performed using PLINK.

Association analysis

Discovery genome-wide meta-analysis

A GWAS of IPF susceptibility was run in each of the Chicago, Colorado and UK studies separately. Analyses were performed using a logistic regression model, assuming an additive genetic effect and adjusting for the first 10 principal components.

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Results were corrected for inflation due to residual fine-scale population structure using genomic control at both the study and meta-analysis level. Genomic control was applied for each study in the discovery meta-analysis ($\lambda = 1.027$ in UK, $\lambda = 1.065$ in Colorado and $\lambda = 1.030$ in Chicago).

These results were combined using a fixed-effect, inverse-variance weighted meta-analysis. Genomic control was also applied to the meta-analysis of all three studies ($\lambda = 1.016$).

Conditional analyses

To identify additional independent signals within each locus in the discovery meta-analysis, conditional analyses were performed by repeating the association analyses for all variants within 1Mb of the sentinel variant, adjusting for the sentinel variant in each study separately and then meta-analysing the results. Variants reaching genome-wide significance after conditioning on the top variant were deemed as independent signals and analyses were repeated until no more independent signals in the region were identified.

Replication analysis

Novel association signals were further tested in the UUS and Genentech studies using a logistic regression model, assuming an additive genetic effect and adjusting for the first 10 principal components. The results were meta-analysed across the UUS and Genentech studies using a fixed-effect, inverse-variance weighted meta-analysis.

Bayesian fine-mapping

Credible sets were calculated for each novel signal to produce a set of variants likely to contain the causal variant at 95% confidence (under the assumption there is a single causal variant and that variant had been measured). Posterior probabilities of the variant being causal were calculated for all variants within 1Mb of the sentinel variant and in at least weak LD with the sentinel variant

($r^2 > 0.1$) in the discovery meta-analysis. Posterior probabilities were calculated from approximate Bayes factors (ABFs) using the formula proposed by Wakefield[6]:

$$\mathrm{ABF} = \frac{1}{\sqrt{1 - \frac{W}{V+W}}} \exp\left(- \frac{Z^2 \quad W}{2\,V+W} \right)$$

where W is the Wakefield prior (which we set to 0.4 which is equivalent to a 95% belief that departure from the null model for the relative risk is less than 1.5), Z is the Z statistic for the variant and V is the variance of the effect size.

The approximate posterior probability was set to equal the ABF for that variant divided by the sum of ABFs for all variants in the signal. Variants were added to the credible set until the sum of the posterior probabilities was greater than or equal to 0.95.

Identification of genes implicated by the association signals

VEP

All variants in each credible set were annotated using VEP[13]. Variants were defined as deleterious if they were recorded as either “deleterious” in SIFT, “probably damaging” in PolyPhen, “likely deleterious” from the CADD score, “likely disease causing” in REVEL, “damaging” in MetaLR or “high” in MutationAssessor.

Association of IPF susceptibility variants with gene expression

Linked genotype and gene expression data resources were interrogated to identify the genes implicated by the novel association signals. Variants in the 95% credible sets were investigated in three eQTL databases; a lung eQTL database consisting of individuals from three cohorts (Universities of British Columbia, Laval and Groningen, n=1,111)[7-9], the NESDA-NTR (Netherlands Study of Depression and Anxiety and the Netherlands Twin Register) blood eQTL database

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

(n=4,896)[10] and 48 tissue types in GTEx (Genotype-Tissue Expression project, n between 80 and 491)[11]. An FDR threshold of 10% was used for the lung eQTL database and NESDA-NTR, and an FDR threshold of 5% was used for the smaller GTEx resource.

Identification of shared causal variants for IPF susceptibility and gene expression changes (colocalisation)

Where IPF susceptibility variants were found to be associated with expression levels of a gene, we tested whether the same variant was likely to be causal both for differences in gene expression and IPF susceptibility. Analyses were performed using the coloc[12] package in R v3.5.1 on all variants in the region with $P < 0.01$ in either the IPF GWAS analysis or eQTL analysis.

The coloc package implements the colocalisation approach described by Giambartolomei et al[12]. In summary, it uses approximate Bayes factors to estimate the probability of each of the following models:

- H_o: There is no association in the region with either IPF risk or the eQTL result
- H_1: There is an association in the region with IPF but not with the expression of the gene
- H_2: There is an association in the region with the expression of the gene but not with IPF
- H_3: There is an association in the region with both IPF and the expression of the gene but these are driven by two different variants
- H_4: There is an association in the region with both IPF and the expression of the gene which is driven by the same variant.

We took colocalisation to be when the probability of H_4 (i.e. the same variant drives IPF risk and the expression of the gene) was greater than 80%.

In silico analyses of functional effects

DeepSEA

DeepSEA[15] (deep learning-based sequence analyzer) is a deep learning method to predict the functional chromatin effects of individual variants. The variant with the highest posterior probability in each of the credible sets for the 14 IPF risk signals identified by the discovery meta-analysis was included in the DeepSEA analysis.

We reported functional effects for any chromatin feature and lung-related cell line that had an E-value < 0.05 (i.e. the expected proportion of SNPs with larger predicted effect for this chromatin feature based on empirical distributions of predicted effects for 1000 Genomes SNPs) and an absolute difference in probability of > 0.1 (threshold for "high confidence") between the reference and alternative allele.

FORGE

FORGE[16] (Functional element overlap analysis of the results of GWAS experiments) is a tool for identifying whether signals in a GWAS are enriched in DNase I hypersensitivity sites in specific tissues. The variant with the highest posterior probability in each of the credible sets for the 14 IPF risk signals identified by the discovery meta-analysis was included in the FORGE analysis. Enrichment was tested in 299 cell lines across 24 tissues including lung and foetal lung.

GARFIELD

GARFIELD[17] (GWAS analysis of regulatory and functional information enrichment with LD correction) is an analysis tool to test if GWAS signals are enriched in functional features. Variants meeting a *P* threshold in the IPF discovery genome-wide analysis were tested for enrichment (*P* thresholds of 5×10^{-8} and 5×10^{-5} were used). Enrichment was tested in DNase I hypersensitivity sites in 424 tissues.

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

SNPsea

SNPsea[18] is a method to identify if gene expression is altered by a set of variants in different cell types or pathways. For this analysis, the variants with the highest posterior probability in each of the credible sets for the 14 IPF risk loci were entered. Genes are implicated using an LD matrix and the expression of these genes are investigated. A score based on the expression of these genes is calculated and compared to a score generated when selecting a random set of variants as the input.

The IPF risk loci were tested for enrichment in genes that showed differential expression between IPF cases and controls in lung epithelium. Expression of genes in four epithelial cell types (normal AT2 cells, indeterminate cells, basal and club/goblet cells) was calculated from lung tissue from six IPF cases and three healthy controls using single cell RNA sequencing data from Xu et al[19]. Gene expression was deemed to be enriched in tissues or pathways if they met a Bonferroni corrected *P* threshold.

Shared genetic susceptibility of IPF, lung function and interstitial lung abnormalities (ILA)

The variant with the highest posterior probability of causality in the credible set for each IPF risk signal were tested for their association with interstitial lung abnormalities (ILA) and lung function.

A genome-wide association analysis of ILAs was conducted by meta-analysis of results from the AGES, COPDGene NHW, ECLIPSE, Framingham, MESA white and SPIROMICS studies. Two analyses were performed; firstly defining cases as any individual with any ILA (n = 1,699) and controls as any individual without an ILA (n = 10,247), and secondly defining cases as those with a subpleural subtype of ILA (n = 1,287) and controls as individuals without any ILA (n = 10,247).

Association with lung function was assessed using data from a genome-wide association study meta-analysis of lung function for 400,102 individuals of European ancestry in UK Biobank and the SpiroMeta consortium[20]. The measures of lung function analysed were FEV_1 (a measure of how

much air an individual can forcibly exhale in the first second), FVC (forced vital capacity, i.e. the total volume of air forcibly exhaled), the ratio of FEV_1/FVC (a measure used in the diagnosis of chronic obstructive pulmonary disease) and PEF (peak expiratory flow, i.e. the highest airflow), which were all measured through spirometry.

Variants were reported as associated with lung function or ILA if they met a Bonferroni corrected *P* value threshold for the number of variants and traits investigated.

Polygenic risk scores

Polygenic risk scores were utilised to assess the contribution of as-yet unreported variants to IPF risk. Polygenic risk scores allow for the cumulative effect of many genetic variants to be studied. The polygenic risk score was equal to the number of risk alleles carried multiplied by the effect size of the variant, summed across all variants included in the score, i.e.:

$$\text{Polygenic Risk Score}_j = \sum_{i=1}^{n} \beta_i X_{ij}$$

where β_i is the log(OR) of variant *i* from the genome-wide meta-analysis of the UK, Chicago and Colorado studies, X_{ij} is the genotype of variant *i* for person *j* and *n* is the number of variants.

Scores were generated for individuals in the independent UUS study using independent variants selected after LD-clumping ($r^2 \leq 0.1$). This score was tested to identify whether it was associated with IPF susceptibility, adjusting for 10 principal components to account for fine-scale population structure, using PRSice v1.25[21]. As we wanted to explore the contribution to IPF risk from variants not yet reported, we excluded variants within 1Mb of each IPF risk locus identified in this IPF susceptibility GWAS. We altered the number of variants included in the risk score calculation by setting a *P*-threshold (P_T) criteria such that the variant had to have a *P* value$<P_T$ in the genome-wide

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

meta-analysis to be included in the score. Given multiple testing, we used the recommended significance threshold of $P<0.001$ for determining significantly associated risk scores[21].

Supplementary Tables

Table E1 - Study level results from discovery genome-wide analysis for novel genome-wide significant and previously reported IPF susceptibility variants

Odds ratios are presented treating the minor allele as the effect allele. Minor allele frequency (MAF) was the allele frequency across the three studies and info is the imputation quality in that study. Results for the sentinel (i.e. most strongly associated) variant from the discovery GWAS in this study are presented except for iii) and iv) where results for the previously reported sentinel variant are shown as there was no association signal observed in the discovery GWAS.

Chr	Pos	Sentinel rsid	r^2 with previous reported sentinel	Locus	Major allele	Minor allele	MAF	Chicago			Colorado			UK			Discovery meta-analysis	
								Info	OR [95% CI]	*P*	Info	OR [95% CI]	*P*	Info	OR [95% CI]	*P*	OR [95% CI]	*P*
i) Novel signals meeting significance criteria																		
3	44902386	rs78238620	-	*KIF15*	T	A	5.3%	0.97	1.74 [1.16, 2.60]	0.007	0.98	1.47 [1.23, 1.78]	4.01×10^{-5}	0.99	1.77 [1.35, 2.33]	4.54×10^{-5}	1.58 [1.37, 1.83]	5.12×10^{-10}
7	1909479	rs12699415	-	*MAD1L1*	G	A	42.0%	0.97	1.43 [1.20, 1.69]	5.31×10^{-5}	0.98	1.23 [1.12, 1.33]	3.67×10^{-6}	0.99	1.30 [1.15, 1.47]	3.51×10^{-5}	1.28 [1.19, 1.37]	7.15×10^{-13}
8	120934126	rs28513081	-	*DEPTOR*	A	G	42.8%	0.99	0.78 [0.66, 0.93]	0.005	0.99	0.84 [0.77, 0.91]	4.69×10^{-5}	0.99	0.79 [0.70, 0.89]	1.94×10^{-4}	0.82 [0.76, 0.87]	1.20×10^{-9}
10	93271016	rs537322302	-	*HECTD2*	C	G	0.3%	-	-	-	0.55	7.52 [2.46, 23.0]	3.98×10^{-4}	0.90	8.04 [3.10, 20.8]	1.79×10^{-5}	7.82 [3.77, 16.2]	3.43×10^{-8}
20	62324391	rs41308092	-	*RTEL1*	G	A	2.1%	0.57	2.06 [1.03, 4.12]	0.040	0.79	2.10 [1.54, 2.86]	2.61×10^{-6}	0.94	2.18 [1.41, 3.37]	4.86×10^{-4}	2.12 [1.67, 2.69]	7.65×10^{-10}
ii) Previously reported signals that reached genome-wide significance in discovery analysis																		
3	169481271	rs12696304	0.97	*LRRC34 /TERC*	C	G	27.9%	0.99	1.38 [1.15, 1.66]	5.57×10^{-4}	0.98	1.33 [1.21, 1.46]	5.52×10^{-9}	1.00	1.22 [1.06, 1.41]	0.005	1.31 [1.21, 1.40]	7.09×10^{-13}
4	89885086	rs2013701	0.28	*FAM13A*	G	T	48.7%	1.00	0.94 [0.79, 1.12]	0.496	1.00	0.78 [0.72, 0.85]	9.16×10^{-9}	1.00	0.72 [0.63, 0.81]	2.27×10^{-7}	0.78 [0.74, 0.84]	3.30×10^{-13}
5	1282414	rs7725218 [a]	0.55	*TERT*	G	A	32.5%	0.52	0.83 [0.69, 1.00]	0.051	0.90	0.68 [0.62, 0.74]	4.88×10^{-17}	1.00	0.76 [0.66, 0.86]	2.68×10^{-5}	0.72 [0.67, 0.77]	1.54×10^{-20}
6	7563232	rs2076295	Same	*DSP*	T	G	46.9%	0.98	1.19 [1.00, 1.42]	0.044	1.00	1.45 [1.33, 1.58]	9.56×10^{-18}	0.99	1.66 [1.47, 1.87]	8.81×10^{-16}	1.46 [1.37, 1.56]	2.79×10^{-30}

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

7	99630342	rs2897075	0.72	7q22.1	C	T	39.1%	0.98	1.30 [1.09, 1.54]	0.003	0.99	1.34 [1.23, 1.46]	2.14×10^{-11}	1.00	1.19 [1.05, 1.36]	0.008	1.30 [1.21, 1.38]	3.10×10^{-14}
11	1241221	rs35705950	Same	*MUC5B*	G	T	14.9%	-	-	-	0.77	4.51 [3.99, 5.09]	1.14×10^{-128}	0.92	5.64 [4.72, 6.73]	3.99×10^{-81}	4.84 [4.37, 5.36]	1.18×10^{-203}
13	113534984	rs9577395	0.86	*ATP11A*	C	G	20.7%	0.77	0.75 [0.61, 0.93]	0.008	0.99	0.75 [0.68, 0.84]	9.27×10^{-8}	1.00	0.81 [0.69, 0.95]	0.008	0.77 [0.71, 0.83]	1.34×10^{-10}
15	40720542	rs59424629	0.97	*IVD*	T	G	46.1%	0.98	0.64 [0.54, 0.76]	5.39×10^{-7}	0.99	0.78 [0.71, 0.85]	3.19×10^{-9}	1.00	0.81 [0.71, 0.92]	9.85×10^{-4}	0.77 [0.71, 0.82]	7.30×10^{-16}
15	86097216	rs62023891	0.54	*AKAP13*	G	A	30.0%	0.97	1.27 [1.06, 1.53]	0.011	0.99	1.25 [1.14, 1.37]	2.98×10^{-5}	0.99	1.30 [1.13, 1.49]	1.86×10^{-4}	1.27 [1.18, 1.36]	1.27×10^{-10}
17	44214888	rs2077551	0.90	*MAPT*	T	C	18.6%	0.85	0.63 [0.51, 0.79]	4.50×10^{-5}	0.86	0.71 [0.64, 0.79]	3.78×10^{-10}	0.96	0.75 [0.65, 0.87]	2.01×10^{-4}	0.71 [0.65, 0.77]	2.83×10^{-16}
19	4717672	rs12610495	Same	*DPP9*	A	G	30.5%	-	-	-	1.00	1.29 [1.17, 1.41]	7.84×10^{-8}	0.97	1.37 [1.20, 1.57]	3.91×10^{-6}	1.31 [1.22, 1.42]	2.92×10^{-12}
iii) Previously reported signals that do not reach genome-wide significance in the discovery meta-analysis																		
6	31864547	rs7887	-	*EHMT2*	G	T	33.9%	0.96	0.92 [0.77, 1.11]	0.388	0.99	0.84 [0.77, 0.92]	1.10×10^{-4}	1.00	1.03 [0.91, 1.18]	0.625	0.90 [0.84, 0.96]	0.002
10	105672842	rs11191865	-	*OBFC1*	G	A	49.1%	1.00	0.98 [0.82, 1.16]	0.809	1.00	1.27 [1.16, 1.38]	5.08×10^{-8}	0.99	1.05 [0.93, 1.19]	0.455	1.16 [1.09, 1.24]	8.91×10^{-6}
14	48040375	rs7144383	-	*MDGA2*	A	G	11.2%	0.98	1.81 [1.37, 2.38]	2.94×10^{-5}	0.88	1.03 [0.90, 1.18]	0.671	0.98	0.95 [0.78, 1.15]	0.568	1.09 [0.98, 1.21]	0.119
iv) Previously reported signals in the 11p15.5 region after conditioning on rs35705950 [b]																		
11	1093945	rs7934606	-	*MUC2*	C	T	44.9%	-	-	-	1.00	0.93 [0.85, 1.02]	0.109	0.94	1.00 [0.87, 1.16]	0.956	0.95 [0.88, 1.03]	0.189
11	1312706	rs111521887	-	*TOLLIP*	C	G	19.8%	-	-	-	0.99	1.00 [0.89, 1.12]	0.972	0.99	1.00 [0.85, 1.19]	0.965	1.00 [0.91, 1.10]	0.996
11	1325829	rs5743890	-	*TOLLIP*	T	C	13.8%	-	-	-	0.93	0.84 [0.74, 0.95]	0.006	0.95	0.88 [0.72, 1.07]	0.193	0.85 [0.76, 0.95]	0.002

[a] This variant was the most significant variant for this signal in the discovery meta-analysis. Although this variant did not quite reach nominal significance ($p<0.05$) in the Chicago study other variants in the signal did reach nominal in each study, had consistent direction of effects in each study and were genome-wide significant in the discovery meta-analysis

[b] The *MUC5B* promoter polymorphism rs35705950 was not imputed in the Chicago study so it was not possible to perform the conditional analysis

Table E2 - Summary of eQTL analysis for novel IPF susceptibility signals

The table below contains all of the genes for which at least one of the variants in the credible set was recorded as an eQTL variant and the tissue this was recorded in. This table also includes the colocalisation probability and IPF risk signals that colocalise with the expression of a gene (taken to be probability > 80%) are shown in green.

Chromosome	GWAS sentinel (risk allele)	eQTL gene	Source	eQTL Tissue (and probe if multiple probes used)	Risk allele effect on gene expression	eQTL sentinel	Colocalisation probability
3	rs78238620_A	*KIF15*	GTEx	Brain - Putamen	Decrease	rs149000267	95.6%
		TMEM42	GTEx	Thyroid	Decrease	rs80059929	93.1%
		KIAA1143	GTEx	Adipose - Subcutaneous	Decrease	rs6792299	0.0%
7	rs12699415_A	*MAD1L1*	NESDA-NTR [a]	Whole blood (11724206_a_at)	Increase	rs35578480	0.1%
				Whole blood (11724207_x_at)	Increase	rs35578480	0.1%
				Whole blood (11737802_a_at)	Increase	rs35578480	0.0%
			GTEx	Heart - Atrial Appendage	Decrease	rs57193072	95.3%
				Nerve - Tibial	Increase	rs7803147	36.3%
		FTSJ2	GTEx	Adipose - Visceral	Decrease	rs34418140	0.0%
				Artery - Aorta	Decrease	rs73673559	0.1%
				Artery - Tibial	Decrease	rs57431109	0.0%
				Brain - Cerebellum	Decrease	rs34418140	0.0%
				Esophagus - Muscularis	Decrease	rs7810970	0.0%
				Muscle - Skeletal	Decrease	rs4719462	0.0%
				Testis	Decrease	rs7810970	0.0%
		AC110781.3	GTEx	Brain - Frontal Cortex	Increase	rs6952808	64.8%
				Brain - Nucleus accumbens	Increase	rs4236272	51.9%
				Testis	Increase	rs10237989	60.5%
8	rs28513081_A	*DEPTOR*	NESDA-NTR [a]	Whole Blood (11751331_a_at)	Increase	rs55892034	93.7%
			Lung eQTL [b]	Lung (100154484_TGI_at)	Decrease	rs1519812	89.5%
				Lung (100312124_TGI_at)	Decrease	rs1519812	89.9%
			GTEx	Brain - Spinal Cord	Decrease	rs72673678	55.2%
				Colon - Sigmoid	Decrease	rs10217077	89.6%
				Colon - Transverse	Decrease	rs7005380	58.1%
				Esophagus - Gastroesophageal Junction	Decrease	rs56177421	0.9%
				Esophagus - Mucosa	Decrease	rs56177421	41.6%
				Esophagus - Muscularis	Decrease	rs56177421	0.9%
				Lung	Decrease	rs10217077	89.2%
				Muscle - Skeletal	Decrease	rs7818296	8.7%

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

				Skin - Not Sun Exposed	Decrease	rs7814520	90.0%
				Skin - Sun Exposed	Decrease	rs1467044	86.5%
		DSCC1	Lung eQTL	Lung (100129753_TGI_at)	Increase	rs7815122	0.0%
				Lung (100132546_TGI_at)	Increase	rs7815122	0.0%
				Lung (100301732_TGI_at)	Increase	rs55741337	0.0%
			GTEx	Artery - Tibial	Increase	rs113408398	0.0%
				Muscle - Skeletal	Increase	rs77647593	0.0%
				Skin - Sun Exposed	Increase	rs28700049	0.0%
		RP11-760H22.2	GTEx	Adipose - Subcutaneous	Increase	rs10217077	84.9%
				Brain - Spinal cord	Decrease	rs7825920	46.8%
				Colon - Sigmoid	Decrease	rs6469868	88.6%
				Esophagus - Gastroesophageal Junction	Decrease	rs56177421	1.0%
				Esophagus - Muscularis	Decrease	rs12541326	0.9%
				Lung	Decrease	rs796666096	90.0%
		KB-1471A8.1	GTEx	Adipose - Subcutaneous	Increase	rs1467044	85.6%
				Adipose - Visceral	Increase	rs7818471	90.9%
				Muscle - Skeletal	Increase	rs7840728	0.1%
				Nerve - Tibial	Increase	rs4870988	3.2%
				Skin - Sun Exposed	Increase	rs13263296	88.7%
				Thyroid	Increase	rs73703111	0.1%
		TAF2	GTEx	Colon - Transverse	Increase	rs112349158	87.5%
20	rs41308092_A	*LIME1*	GTEx	Muscle - Skeletal	Increase	rs4809330	20.5%

[a] Only results for significantly associated variants in the NESDA-NTR dataset were available, therefore the colocalisation analysis was run only including variants significantly associated with gene expression in blood rather than all variants in the region.

[b] The lung eQTL dataset showed two independent signals of association for *DEPTOR* expression. The eQTL results here are those obtained after conditioning on the top eQTL for *DEPTOR* to condition out the strongest signal which was driven by different variants to those driving the IPF risk association.

Table E3 - Annotation and eQTL results for variants in 95% credible sets of novel IPF susceptibility signals

i) Chromosome 3

rsid	chr	Position	GWAS *P*	Posterior Probability	Annotation	R^2 with sentinel	eQTL			
							Lung eQTL	GTEx (lung)	GTEx (non-lung tissue)	NESDA-NTR
rs78238620	3	44902386	5.12×10^{-10}	14.71%	intron (*KIF15*)	Sentinel	-	-	*KIF15, TMEM42, KIAA1143*	-
rs2292180	3	44903349	5.42×10^{-10}	14.01%	intron (*KIF15*)	1.00	-	-	*KIF15, TMEM42, KIAA1143*	-
rs2292181	3	44903434	5.42×10^{-10}	14.01%	synonymous (*TMEM42*), intron (*KIF15*), non-coding exon (*MIR564*)	1.00	-	-	*KIF15, TMEM42, KIAA1143*	-
rs74341405	3	44845649	7.74×10^{-10}	8.90%	intron (*KIF15*)	0.89	-	-	*KIF15, TMEM42*	-
rs80059929	3	44846722	7.74×10^{-10}	8.90%	intron (*KIF15*)	0.89	-	-	*KIF15, TMEM42*	-
rs76304484	3	44877209	2.20×10^{-9}	4.24%	intron (*KIF15*)	0.99	-	-	*KIF15, TMEM42, KIAA1143*	-
rs6792918	3	44857004	2.83×10^{-9}	3.43%	intron (*KIF15*)	0.99	-	-	*KIF15, TMEM42, KIAA1143*	-
rs77568017	3	44877853	3.57×10^{-9}	2.79%	intron (*KIF15*)	0.99		-	*KIF15, TMEM42, KIAA1143*	-
rs76526953	3	44881909	3.57×10^{-9}	2.79%	intron (*KIF15*)	0.99	-	-	*KIF15, TMEM42, KIAA1143*	-
rs141979279	3	44858131	4.01×10^{-9}	2.52%	intron (*KIF15*)	0.99	-	-	*KIF15, TMEM42, KIAA1143*	-
rs55661644	3	44869509	4.01×10^{-9}	2.52%	intron (*KIF15*)	0.99	-	-	*KIF15, TMEM42, KIAA1143*	-
rs7340559	3	44871986	4.01×10^{-9}	2.52%	intron (*KIF15*)	0.99	-	-	*KIF15, TMEM42, KIAA1143*	-
rs77136835	3	44874693	4.01×10^{-9}	2.52%	intron (*KIF15*)	0.99	-	-	*KIF15, TMEM42, KIAA1143*	-
rs112645395	3	44794881	3.97×10^{-9}	1.92%	missense (*KIAA1143*) [SIFT: Deleterious PolyPhen: Possibly damaging CADD: Likely benign REVEL: Likely benign MetaLR: Tolerated MutationAssessor: Medium]	0.79	-	-	*TMEM42*	-
rs149000267	3	44836326	3.09×10^{-9}	1.56%	intron (*KIF15*)	0.64	-	-	*KIF15, TMEM42*	-
rs77938604	3	44836543	7.81×10^{-9}	1.28%	intron (*KIF15*)	0.90	-	-	*KIF15, KIAA1143*	-
rs4682996	3	44819436	1.05×10^{-8}	0.99%	intron (*KIF15*)	0.89	-	-	*KIAA1143, TMEM42*	-
rs4682992	3	44786946	1.22×10^{-8}	0.88%	intron (*KIAA1143*)	0.89	-	-	*KIAA1143, TMEM42*	-
rs112842175	3	44788306	1.22×10^{-8}	0.88%	intron (*KIAA1143*)	0.89	-	-	*KIAA1143, TMEM42*	-

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

rs4682993	3	44795238	1.22×10^{-8}	0.88%	intron (*KIAA1143*)	0.89	-	-	*KIAA1143, TMEM42*	-
rs77805183	3	44797277	1.22×10^{-8}	0.88%	intron (*KIAA1143*)	0.89	-	-	*KIAA1143, TMEM42*	-
rs111788055	3	44833973	1.36×10^{-8}	0.81%	intron (*KIF15*)	0.90	-	-	*KIF15, KIAA1143*	-
rs79850585	3	44756245	1.38×10^{-8}	0.80%	intron (*ZNF502*)	0.88	-	-	*KIAA1143, TMEM42*	-
rs4682994	3	44803130	1.58×10^{-8}	0.71%	intron (*KIF15*)	0.89	-	-	*KIF15, TMEM42, KIAA1143*	-

ii) Chromosome 7

rsid	Chr	Position	GWAS *P*	Posterior Probability	Annotation	R^2 with Sentinel	eQTL			
							Lung eQTL	GTEx (lung)	GTEx (non-lung tissue)	NESDA-NTR
rs12699415	7	1909479	7.15×10^{-13}	35.67%	intron (*MAD1L1*)	Sentinel	-	-	*MAD1L1*	-
rs7795126	7	2076626	8.47×10^{-12}	3.38%	intron (*MAD1L1*)	0.71	-	-	*MAD1L1*	-
rs10950503	7	2039594	1.10×10^{-11}	2.64%	intron (*MAD1L1*)	0.80	-	-	*MAD1L1*	-
rs4455739	7	1864356	1.12×10^{-11}	2.61%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs34373690	7	1869473	1.27×10^{-11}	2.31%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs34074471	7	1865249	1.27×10^{-11}	2.30%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs35091011	7	1865174	1.33×10^{-11}	2.21%	intron (*MAD1L1*)	0.68		-	*AC110781.3*	-
rs35406566	7	1865527	1.33×10^{-11}	2.21%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs6974455	7	1865921	1.35×10^{-11}	2.17%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs61164094	7	1874264	1.37×10^{-11}	2.14%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs4255035	7	1864444	1.41×10^{-11}	2.09%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs4379359	7	1864415	1.52×10^{-11}	1.94%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs7806394	7	1864129	1.53×10^{-11}	1.93%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs4631355	7	1864245	1.53×10^{-11}	1.93%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs3857706	7	2034193	1.53×10^{-11}	1.93%	intron (*MAD1L1*)	0.81	-	-	*MAD1L1*	-
rs13225346	7	1866916	1.59×10^{-11}	1.86%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs57193069	7	1862417	1.74×10^{-11}	1.71%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs872464	7	2034562	1.93×10^{-11}	1.54%	intron (*MAD1L1*)	0.81	-	-	*MAD1L1*	-
rs6955652	7	1865583	1.98×10^{-11}	1.51%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs7799807	7	1868092	2.01×10^{-11}	1.48%	intron (*MAD1L1*)	0.37	-	-	*AC110781.3, FTSJ2*	-
rs35641411	7	1870242	2.03×10^{-11}	1.48%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs1403174	7	2032865	2.03×10^{-11}	1.47%	intron (*MAD1L1*)	0.82	-	-	*MAD1L1*	-
rs28661143	7	1866395	2.24×10^{-11}	1.34%	intron (*MAD1L1*)	0.68	-	-	*AC110781.3*	-
rs12537430	7	1868761	3.25×10^{-11}	0.93%	intron (*MAD1L1*)	0.36	-	-	*AC110781.3, FTSJ2*	-
rs12537479	7	1868995	3.48×10^{-11}	0.88%	intron (*MAD1L1*)	0.36	-	-	*AC110781.3, FTSJ2*	-
rs35935754	7	1869242	3.60×10^{-11}	0.85%	intron (*MAD1L1*)	0.36	-	-	*AC110781.3, FTSJ2*	-

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

rs7799782	7	1868039	4.05×10^{-11}	0.76%	intron (*MAD1L1*)	0.36	-	-	*AC110781.3, FTSJ2*	-
rs6959688	7	1966831	4.33×10^{-11}	0.72%	intron (*MAD1L1*)	0.86	-	-	*MAD1L1*	-
rs56053419	7	1863463	5.32×10^{-11}	0.59%	intron (*MAD1L1*)	0.36	-	-	*AC110781.3, FTSJ2*	-
rs55948146	7	1866953	8.08×10^{-11}	0.39%	intron (*MAD1L1*)	0.36	-	-	*AC110781.3, FTSJ2*	-
rs12672286	7	1907009	8.29×10^{-11}	0.38%	intron (*MAD1L1*)	0.38	-	-	*AC110781.3, FTSJ2*	-
rs6978112	7	1966841	9.52×10^{-11}	0.34%	intron (*MAD1L1*)	0.84	-	-	*MAD1L1*	-
rs4721090	7	1873084	1.03×10^{-10}	0.31%	intron (*MAD1L1*)	0.32	-	-	*AC110781.3, FTSJ2*	-
rs11761670	7	1904709	1.12×10^{-10}	0.29%	intron (*MAD1L1*)	0.37	-	-	*AC110781.3, FTSJ2*	-
rs13224015	7	1913869	1.11×10^{-10}	0.29%	intron (*MAD1L1*)	0.84	-	-	*MAD1L1*	*MAD1L1*
rs4721143	7	1918179	1.14×10^{-10}	0.28%	intron (*MAD1L1*)	0.84	-	-	*MAD1L1*	*MAD1L1*
rs34120092	7	1861952	1.25×10^{-10}	0.26%	intron (*MAD1L1*)	0.35	-	-	*AC110781.3, FTSJ2*	-
rs7786367	7	1863828	1.27×10^{-10}	0.26%	intron (*MAD1L1*)	0.71	-	-	*AC110781.3*	-
rs13235380	7	1873894	1.27×10^{-10}	0.26%	intron (*MAD1L1*)	0.32	-	-	*AC110781.3, FTSJ2*	-
rs9770241	7	1863164	1.37×10^{-10}	0.24%	intron (*MAD1L1*)	0.37	-	-	*AC110781.3, FTSJ2*	-
rs10807751	7	1883476	1.39×10^{-10}	0.24%	intron (*MAD1L1*), intron (*AC110781*.3)	0.70	-	-	*AC110781.3*	-
rs13221208	7	1913856	1.35×10^{-10}	0.24%	intron (*MAD1L1*)	0.84	-	-	*MAD1L1*	*MAD1L1*
rs34256344	7	1861460	1.44×10^{-10}	0.23%	intron (*MAD1L1*)	0.72	-	-	*AC110781.3*	-
rs10237989	7	1873343	1.42×10^{-10}	0.23%	intron (*MAD1L1*)	0.32	-	-	*AC110781.3, FTSJ2*	-
rs13222183	7	1873879	1.44×10^{-10}	0.23%	intron (*MAD1L1*)	0.32	-	-	*AC110781.3, FTSJ2*	-
rs6460944	7	1876199	1.45×10^{-10}	0.23%	intron (*MAD1L1*)	0.58	-	-	*AC110781.3*	*MAD1L1*
rs12537387	7	1868582	1.64×10^{-10}	0.20%	intron (*MAD1L1*)	0.57	-	-	-	*MAD1L1*
rs4719319	7	1888094	1.68×10^{-10}	0.20%	intron (*MAD1L1*)	0.33	-	-	*AC110781.3, FTSJ2*	-
rs4719330	7	1914613	1.65×10^{-10}	0.20%	intron (*MAD1L1*)	0.85	-	-	*MAD1L1*	*MAD1L1*
rs4721139	7	1917337	1.64×10^{-10}	0.20%	intron (*MAD1L1*)	0.84	-	-	*MAD1L1*	*MAD1L1*
rs6949794	7	1908727	1.83×10^{-10}	0.18%	intron (*MAD1L1*)	0.39	-	-	*MAD1L1, AC110781.3, FTSJ2*	-
rs10950400	7	1882470	1.99×10^{-10}	0.17%	intron (*MAD1L1*), intron (*AC110781*.3)	0.33	-	-	*AC110781.3, FTSJ2*	-
rs4639400	7	1917806	1.93×10^{-10}	0.17%	intron (*MAD1L1*)	0.84	-	-	*MAD1L1*	*MAD1L1*

rs7783715	7	1923385	2.00×10^{-10}	0.17%	intron (*MAD1L1*)	0.85	-	-	-	*MAD1L1*
rs6977733	7	1886725	2.05×10^{-10}	0.16%	intron (*MAD1L1*), intron (*AC110781*.3)	0.33	-	-	*AC110781.3, FTSJ2*	-
rs6954521	7	1886865	2.13×10^{-10}	0.16%	intron (*MAD1L1*), intron (*AC110781*.3)	0.33	-	-	*AC110781.3, FTSJ2*	-
rs6978048	7	1886872	2.13×10^{-10}	0.16%	intron (*MAD1L1*), intron (*AC110781*.3)	0.33	-	-	*AC110781.3, FTSJ2*	-
rs6954673	7	1886937	2.05×10^{-10}	0.16%	intron (*MAD1L1*), missense (*AC110781*.3) [PolyPhen: Benign CADD: Likely benign]	0.33	-	-	*AC110781.3, FTSJ2*	-
rs4610628	7	1903100	2.08×10^{-10}	0.16%	intron (*MAD1L1*)	0.36	-	-	*AC110781.3, FTSJ2*	-
rs10950411	7	1909153	2.13×10^{-10}	0.16%	intron (*MAD1L1*)	0.85	-	-	*MAD1L1*	*MAD1L1*
rs4458759	7	1876081	2.23×10^{-10}	0.15%	intron (*MAD1L1*)	0.58	-	-	*AC110781.3*	*MAD1L1*
rs6948403	7	1876768	2.28×10^{-10}	0.15%	intron (*MAD1L1*)	0.58	-	-	*AC110781.3*	*MAD1L1*
rs6953693	7	1886388	2.25×10^{-10}	0.15%	intron (*MAD1L1*), intron (*AC110781*.3)	0.33	-	-	*AC110781.3, FTSJ2*	-
rs10950410	7	1909086	2.20×10^{-10}	0.15%	intron (*MAD1L1*)	0.73	-	-	*AC110781.3*	*MAD1L1*
rs57216949	7	2030287	2.30×10^{-10}	0.15%	intron (*MAD1L1*)	0.37	-	-	*MAD1L1, FTSJ2*	-
rs12534763	7	1868711	2.45×10^{-10}	0.14%	intron (*MAD1L1*)	0.58	-	-	-	*MAD1L1*
rs6965935	7	1876895	2.44×10^{-10}	0.14%	intron (*MAD1L1*)	0.58	-	-	*AC110781.3*	*MAD1L1*
rs6957894	7	1887362	2.38×10^{-10}	0.14%	intron (*MAD1L1*), missense (*AC110781*.3) [PolyPhen: Probably damaging CADD: Likely benign]	0.33	-	-	*AC110781.3, FTSJ2*	-
rs4719318	7	1887930	2.38×10^{-10}	0.14%	intron (*MAD1L1*)	0.33	-	-	*AC110781.3, FTSJ2*	-
rs35349665	7	1911166	2.48×10^{-10}	0.14%	intron (*MAD1L1*)	0.84	-	-	-	*MAD1L1*
rs4721134	7	1912057	2.35×10^{-10}	0.14%	intron (*MAD1L1*)	0.39	-	-	*AC110781.3, FTSJ2*	-
rs4449693	7	1884630	2.53×10^{-10}	0.13%	intron (*MAD1L1*), intron (*AC110781*.3)	0.33	-	-	*AC110781.3, FTSJ2*	-
rs6952808	7	1886535	2.66×10^{-10}	0.13%	intron (*MAD1L1*), intron (*AC110781*.3)	0.33	-	-	*AC110781.3, FTSJ2*	-

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

rs10260585	7	1889521	2.54×10^{-10}	0.13%	intron (*MAD1L1*)	0.33	-	-	*AC110781.3, FTSJ2*	-
rs4256490	7	1890764	2.63×10^{-10}	0.13%	intron (*MAD1L1*)	0.33	-	-	*AC110781.3, FTSJ2*	-
rs4601204	7	1890925	2.51×10^{-10}	0.13%	intron (*MAD1L1*)	0.33	-	-	*AC110781.3, FTSJ2*	-
rs6948707	7	1870794	2.89×10^{-10}	0.12%	intron (*MAD1L1*)	0.58	-	-	-	*MAD1L1*
rs3889797	7	1877924	2.91×10^{-10}	0.12%	intron (*MAD1L1*)	0.58	-	-	*AC110781.3*	*MAD1L1*
rs4721122	7	1893311	2.74×10^{-10}	0.12%	intron (*MAD1L1*)	0.33	-	-	*AC110781.3, FTSJ2*	-
rs12155225	7	1899479	2.91×10^{-10}	0.12%	intron (*MAD1L1*)	0.38	-	-	*MAD1L1, AC110781.3, FTSJ2*	-
rs12538674	7	1925166	2.79×10^{-10}	0.12%	intron (*MAD1L1*)	0.85	-	-	*MAD1L1*	*MAD1L1*
rs4721287	7	2028663	2.75×10^{-10}	0.12%	intron (*MAD1L1*)	0.37	-	-	*MAD1L1, FTSJ2*	-
rs56727870	7	2029940	2.98×10^{-10}	0.11%	intron (*MAD1L1*)	0.37	-	-	*MAD1L1, FTSJ2*	-
rs60995052	7	2030007	3.03×10^{-10}	0.11%	intron (*MAD1L1*)	0.37	-	-	*MAD1L1, FTSJ2*	-
rs60755037	7	2030104	3.03×10^{-10}	0.11%	intron (*MAD1L1*)	0.37	-	-	*MAD1L1, FTSJ2*	-
7:2036550	7	2036550	2.97×10^{-10}	0.11%	intron (*MAD1L1*)	0.37	-	-	-	-

iii) Chromosome 8

rsid	Chr	Position	GWAS *P*	Posterior Probability	Annotation	R^2 with Sentinel	eQTL			
							Lung eQTL	GTEx (lung)	GTEx (non-lung tissue)	NESDA-NTR
rs28513081	8	120934126	1.20×10^{-9}	4.51%	intron (*DEPTOR*)	Sentinel	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs6469878	8	120938448	1.66×10^{-9}	3.33%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs7814294	8	120939436	1.71×10^{-9}	3.23%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs2037346	8	120935452	1.73×10^{-9}	3.19%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs10808505	8	120940206	1.85×10^{-9}	3.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs4871787	8	120936418	2.08×10^{-9}	2.68%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs1464276	8	120937041	2.21×10^{-9}	2.53%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs10107579	8	120934569	2.62×10^{-9}	2.16%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs56864850	8	120943444	2.74×10^{-9}	2.07%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs55892034	8	120943507	2.74×10^{-9}	2.07%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs9987332	8	120933963	4.18×10^{-9}	1.39%	intron (*DEPTOR*)	0.96	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs13265546	8	120919975	5.01×10^{-9}	1.17%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs35006524	8	120930769	5.28×10^{-9}	1.11%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs1607624	8	120929289	5.90×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs7829901	8	120929834	5.90×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs13267896	8	120920654	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

rs13275524	8	120920941	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs6469867	8	120921412	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs6469868	8	120921841	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs6469871	8	120922079	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs6469872	8	120922247	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs7015470	8	120922341	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs6995139	8	120922397	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs10217077	8	120923183	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs10217083	8	120923285	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs939242	8	120924270	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs939241	8	120924537	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs10216503	8	120925083	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs11785871	8	120925199	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs7824545	8	120925621	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs7387264	8	120925998	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs7388508	8	120926153	5.91×10^{-9}	1.00%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs7462302	8	120917771	6.15×10^{-9}	0.96%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs12545863	8	120918111	6.15×10^{-9}	0.96%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-

rs7002839	8	120918748	6.15×10^{-9}	0.96%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs7006905	8	120918923	6.15×10^{-9}	0.96%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs9720591	8	120916840	6.36×10^{-9}	0.93%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs4073560	8	120917388	6.36×10^{-9}	0.93%	intron (*DEPTOR*)	0.93	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs10110216	8	120887547	7.01×10^{-9}	0.85%	intron (*DEPTOR*)	0.84	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs10110223	8	120887566	7.05×10^{-9}	0.84%	intron (*DEPTOR*)	0.84	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs13257252	8	120912233	7.13×10^{-9}	0.84%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs35272074	8	120910783	7.39×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs10103660	8	120910787	7.39×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs13281299	8	120909493	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs7814496	8	120909628	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs7814520	8	120909665	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs7832923	8	120909780	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs6469861	8	120909913	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs6469862	8	120910036	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs6469863	8	120910150	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs6469864	8	120910225	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs6469865	8	120910235	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

rs6987580	8	120910380	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs6988011	8	120910538	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs6993375	8	120911630	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs6993797	8	120911645	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs6989741	8	120912154	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs13258592	8	120912429	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs10093230	8	120912846	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs9649947	8	120913357	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs7465609	8	120913860	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs7465612	8	120913885	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs13279398	8	120914270	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs13277947	8	120914432	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs13277992	8	120914527	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs13253140	8	120914908	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs11777705	8	120915232	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs12681402	8	120915862	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs7461290	8	120916394	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs9721029	8	120916555	7.40×10^{-9}	0.81%	intron (DEPTOR)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*

rs9721042	8	120916651	7.40×10^{-9}	0.81%	intron (*DEPTOR*)	0.92	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs10098306	8	120908921	7.50×10^{-9}	0.80%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs7459671	8	120908886	8.66×10^{-9}	0.70%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs7465181	8	120908589	8.72×10^{-9}	0.69%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs7462250	8	120908607	9.14×10^{-9}	0.66%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs1467044	8	120887041	9.71×10^{-9}	0.62%	intron (*DEPTOR*)	0.84	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs7003126	8	120888752	9.71×10^{-9}	0.62%	intron (*DEPTOR*)	0.84	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs11781657	8	120889242	9.71×10^{-9}	0.62%	intron (*DEPTOR*)	0.84	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs4871013	8	120907990	9.73×10^{-9}	0.62%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs4871772	8	120907895	1.21×10^{-8}	0.51%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs4871012	8	120907974	1.21×10^{-8}	0.51%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs4871773	8	120907911	1.21×10^{-8}	0.51%	intron (*DEPTOR*)	0.92	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs13265367	8	120904009	1.74×10^{-8}	0.36%	intron (*DEPTOR*)	0.90	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs10107251	8	120904348	1.74×10^{-8}	0.36%	intron (*DEPTOR*)	0.90	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs10107361	8	120904427	1.74×10^{-8}	0.36%	intron (*DEPTOR*)	0.90	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs10094455	8	120904676	1.74×10^{-8}	0.36%	intron (*DEPTOR*)	0.90	*DEPTOR*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs10094458	8	120904688	1.74×10^{-8}	0.36%	intron (*DEPTOR*)	0.90	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs10094587	8	120904800	1.74×10^{-8}	0.36%	intron (*DEPTOR*)	0.90	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

rs13275184	8	120905104	1.74×10^{-8}	0.36%	intron (*DEPTOR*)	0.90	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	-
rs13249122	8	120905919	1.74×10^{-8}	0.36%	intron (*DEPTOR*)	0.90	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs13250594	8	120906391	1.74×10^{-8}	0.36%	intron (*DEPTOR*)	0.90	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs13259990	8	120907183	1.75×10^{-8}	0.36%	intron (*DEPTOR*)	0.90	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs13261304	8	120907420	1.79×10^{-8}	0.35%	intron (*DEPTOR*)	0.90	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs12681623	8	120903856	1.81×10^{-8}	0.35%	intron (*DEPTOR*)	0.90	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*
rs13260933	8	120907250	1.86×10^{-8}	0.34%	intron (*DEPTOR*)	0.90	*DEPTOR, DSCC1*	*DEPTOR, RP11-760H22.2*	*DEPTOR, DSCC1, KB-1471A8.1, RP11-760H22.2, TAF2*	*DEPTOR*

iv) Chromosome 10

rsid	Chr	Position	GWAS *P*	Posterior Probability	Annotation	R^2 with sentinel	eQTL			
							Lung eQTL	GTEx (lung)	GTEx (non-lung tissue)	NESDA-NTR
rs537322302	10	93271016	3.43×10^{-8}	48.25%	intron (*HECTD2*)	Sentinel	-	-	-	-
rs547164341	10	93285553	3.41×10^{-6}	42.03%	intergenic	0.77	-	-	-	-
rs143984698	10	93059485	0.012	5.31%	intergenic	0.20	-	-	-	-

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

v) Chromosome 20

rsid	chr	Position	GWAS *P*	Posterior Probability	Annotation	R^2 with sentinel	eQTL			
							Lung eQTL	GTEx (lung)	GTEx (non-lung tissue)	NESDA-NTR
rs116905074	20	62377853	8.28×10^{-10}	33.16%	intron (*ZBTB46*)	0.71	-	-	-	-
rs41308092	20	62324391	7.65×10^{-10}	28.99%	intron (*RTEL1*, *RTEL1-TNFRSF6B*)	Sentinel	-	-	*LIME1*	-
rs118130858	20	62345681	1.28×10^{-9}	22.25%	intron (*ZGPAT*, *RP4-583P15.15*)	0.81	-	-	-	-
rs115610405	20	62325833	2.42×10^{-9}	14.20%	missense (*RTEL1*, *RTEL1-TNFRSF6B*) [SIFT: Tolerated PolyPhen: Possibly damaging CADD: Likely benign REVEL: Likely benign MetaLR: Tolerated MutationAssessor: Low]	0.86	-	-	-	-

Table E4 - DeepSEA results for predicted chromatic effects of rs2013701

Predicted chromatin effects from the DeepSEA analysis for the rs2013701 variant in *FAM13A*. Difference is the difference between the probabilities of chromatin effects between the alternative and reference allele. E-value is the expected proportion of SNPs with larger predicted effect (from reference allele to alternative allele) for this chromatin feature computed based on the empirical distributions of predicted effects for 1000 Genomes SNPs. We report predicted functional effects as those with difference > 0.1 and E-value < 0.05.

Chromatin feature	Cell line (treatment)	Fold change	Difference	E-value
DNase	Prostate adenocarcinoma (androgen)	0.840	0.130	7.32×10^{-4}
DNase	Primary tracheal epithelial cells	1.281	0.111	0.003
DNase	Epithelial cell line derived from a mammary ductal carcinoma	0.988	0.116	6.82×10^{-4}
DNase	Epithelial cell line derived from a lung carcinoma tissue	1.086	0.152	0.002
DNase	Mammary epithelial cells	1.534	0.179	0.002
DNase	Prostate adenocarcinoma	0.843	0.134	0.002
DNase	Mammary gland	1.106	0.187	7.89×10^{-4}
DNase	Choroid plexus epithelial cells	1.883	0.112	0.003
DNase	Esophageal epithelial cells	1.662	0.177	0.003
DNase	Iris pigment epithelial cells	1.949	0.146	0.004
DNase	Renal cortical epithelial cells	1.804	0.151	0.002
DNase	Renal epithelial cells	2.011	0.121	0.002
DNase	Villous mesenchymal fibroblast cells	1.660	0.106	0.003
DNase	Pancreatic carcinoma	1.258	0.130	0.001
DNase	Prostate epithelial cell line	1.803	0.254	0.001
DNase	Small airway epithelial cells	1.672	0.194	0.002
DNase	Embryonic lung fibroblast cells (4OHTAM_20nM_72hr)	1.654	0.116	0.003
DNase	Embryonic lung fibroblast cells	1.747	0.122	0.003
FOXA1	Epithelial cell line derived from a mammary ductal carcinoma (DMSO_0.02pct)	0.592	0.102	0.002

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Table E5 - SNPsea results for enrichment of IPF susceptibility signals in IPF specific differentially expressed genes across four lung epithelial cell types

Cell type	*P* value
Normal AT2 cells	0.366
Indeterminate cells	0.485
Basal cells	0.236
Club/Goblet	0.475

Table E6 - Results for IPF risk variants in interstitial lung abnormalities and lung function GWAS

This table shows the results from the interstitial lung abnormality (ILA) and lung function analyses for the 16 signals that reached genome-wide significance in the IPF case-control discovery meta-analysis. All results are presented for the allele that is associated with increased risk of IPF. The IPF odds ratios are the discovery meta-analysis results. FEV_1 is the forced expiratory volume after 1 second, FVC is the forced vital capacity and PEF is the peak expiratory flow. Variants that show an association after multiple testing corrections for 96 tests ($P < 0.00052$) are shown in green (variants that have $P < 0.05$ but do not reach significance are shown in yellow).

Chr	rsid	Locus	Effect allele	EAF	IPF	Any ILA	Subpleural ILA	FEV_1	FVC	FEV_1 / FVC	PEF
					OR [95% CI] *P*	OR [95% CI] *P*	OR [95% CI] *P*	β [95% CI] *P*	β [95% CI] *P*	β [95% CI] *P*	β [95% CI] *P*
3	rs78238620	*KIF15*	A	5.3%	1.54 [1.38, 1.73] 2.94×10^{-14}	1.15 [0.95, 1.40] 0.161	1.18 [0.95, 1.46] 0.126	−0.011 [−0.022, 0.000] 0.069	−0.022 [-0.033, 0.011] 2.92×10^{-4}	0.017 [0.006, 0.028] 0.005	0.000 [−0.011, 0.011] 0.851
3	rs12696304	*LRRC34/TERC*	G	27.9%	1.31 [1.21, 1.40] 7.09×10^{-13}	1.10 [0.99, 1.21] 0.073	1.18 [1.05, 1.31] 0.004	−0.006 [−0.011, 0.000] 0.046	−0.007 [−0.012, −0.001] 0.023	0.001 [−0.004, 0.007] 0.673	0.001 [−0.005, 0.006] 0.770
4	rs2013701	*FAM13A*	G	51.3%	1.28 [1.19, 1.35] 3.30×10^{-13}	1.13 [1.03, 1.23] 0.009	1.11 [1.01, 1.23] 0.039	0.006 [0.001, 0.011] 0.013	−0.014 [−0.019, −0.009] 1.02×10^{-7}	0.043 [0.038, 0.048] 3.56×10^{-60}	0.019 [0.014, 0.024] 2.16×10^{-13}
5	rs7725218	*TERT*	G	67.5%	1.39 [1.30, 1.49] 1.54×10^{-20}	1.05 [0.95, 1.16] 0.333	1.06 [0.95, 1.18] 0.310	−0.005 [−0.010, 0.000] 0.042	−0.007 [−0.012, −0.002] 0.007	0.004 [−0.001, 0.009] 0.154	0.001 [−0.004, 0.007] 0.526
6	rs2076295	*DSP*	G	46.9%	1.46 [1.37, 1.56] 2.79×10^{-30}	1.18 [1.08, 1.29] 2.84×10^{-4}	1.21 [1.10, 1.33] 1.56×10^{-4}	−0.002 [−0.007, 0.003] 0.432	−0.013 [−0.018, −0.009] 4.22×10^{-7}	0.023 [0.019, 0.028] 2.52×10^{-19}	0.006 [0.001, 0.011] 0.020
7	rs12699415	*MAD1L1*	A	42.0%	1.28 [1.22, 1.35] 5.50×10^{-20}	1.04 [0.95, 1.14] 0.438	0.98 [0.89, 1.09] 0.707	−0.007 [−0.012, −0.002] 0.011	−0.011 [−0.016, −0.007] 1.41×10^{-5}	0.008 [0.003, 0.012] 0.005	−0.001 [−0.006, 0.004] 0.581
7	rs2897075	7q22.1	T	39.1%	1.30 [1.21, 1.38] 3.10×10^{-14}	1.03 [0.94, 1.13] 0.560	1.07 [0.97, 1.19] 0.165	0.007 [0.003, 0.012] 0.004	−0.004 [-0.009, 0.000] 0.143	0.025 [0.020, 0.030] 6.06×10^{-21}	0.012 [0.007, 0.017] 2.86×10^{-6}
8	rs28513081	*DEPTOR*	A	57.2%	1.20 [1.14, 1.27] 1.84×10^{-11}	1.03 [0.94, 1.12] 0.494	1.05 [0.95, 1.16] 0.331	0.001 [−0.004, 0.006] 0.822	−0.005 [−0.010, −0.001] 0.045	0.011 [0.006, 0.016] 4.22×10^{-5}	0.005 [0.000, 0.010] 0.073

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

10	rs537322302	*HECTD2*	G	0.3%	3.82 [2.25, 6.48] 7.41×10^{-7}	3.25 [1.01, 10.5] 0.049	6.23 [1.87, 20.8] 0.003	−0.005 [−0.042, 0.032] 0.832	−0.014 [−0.051, 0.023] 0.524	0.019 [−0.018, 0.056] 0.386	0.021 [−0.016, 0.059] 0.312
11	rs35705950	*MUC5B*	T	14.9%	4.84 [4.37, 5.36] 1.18×10^{-203}	1.98 [1.75, 2.24] 4.45×10^{-27}	2.23 [1.94, 2.57] 4.94×10^{-29}	−0.005 [−0.013, 0.002] 0.147	−0.013 [−0.021, −0.006] 8.21×10^{-4}	0.015 [0.008, 0.023] 7.11×10^{-4}	0.008 [0.000, 0.016] 0.063
13	rs9577395	*ATP11A*	C	79.3%	1.30 [1.20, 1.41] 1.34×10^{-10}	1.07 [0.96, 1.19] 0.236	1.08 [0.96, 1.21] 0.225	0.011 [0.005, 0.017] 2.82×10^{-4}	0.006 [0.000, 0.012] 0.041	0.012 [0.007, 0.018] 2.04×10^{-4}	0.011 [0.005, 0.017] 3.70×10^{-4}
15	rs59424629	*IVD*	T	53.8%	1.31 [1.22, 1.40] 7.30×10^{-16}	1.11 [1.01, 1.20] 0.027	1.15 [1.04, 1.25] 0.005	−0.014 [−0.019, −0.009] 1.14×10^{-7}	−0.015 [−0.020, 0.011] 6.97×10^{-9}	0.000 [−0.005, 0.004] 0.851	−0.005 [−0.000, 0.000] 0.072
15	rs62023891	*AKAP13*	A	30.0%	1.27 [1.18, 1.36] 1.27×10^{-10}	1.14 [1.03, 1.25] 0.008	1.09 [0.98, 1.21] 0.124	−0.001 [−0.006, 0.004] 0.726	0.001 [−0.005, 0.006] 0.834	−0.003 [−0.008, 0.003] 0.356	−0.004 [−0.010, 0.001] 0.109
17	rs2077551	*MAPT*	T	81.4%	1.41 [1.30, 1.54] 2.83×10^{-16}	1.23 [1.07, 1.41] 0.003	1.23 [1.05, 1.43] 0.008	0.045 [0.039, 0.052] 1.02×10^{-42}	0.044 [0.038, 0.050] 3.21×10^{-39}	0.012 [0.006, 0.018] 5.81×10^{-4}	0.025 [0.018, 0.031] 2.04×10^{-13}
19	rs12610495	*DPP9*	G	30.5%	1.31 [1.22, 1.42] 2.92×10^{-12}	1.12 [1.00, 1.25] 0.041	1.22 [1.08, 1.37] 0.001	−0.001 [−0.006, 0.004] 0.682	−0.004 [−0.009, 0.001] 0.127	0.004 [−0.001, 0.009] 0.153	0.005 [0.000, 0.011] 0.051
20	rs41308092	*RTEL1*	A	2.1%	1.83 [1.52, 2.20] 1.38×10^{-10}	0.76 [0.48, 1.19] 0.232	0.87 [0.54, 1.40] 0.563	−0.016 [−0.033, 0.001] 0.084	−0.022 [−0.039, −0.004] 0.022	0.009 [−0.008, 0.027] 0.321	−0.011 [−0.028, 0.007] 0.335

Table E7 - Results from IPF discovery meta-genome-wide analysis for the 279 variants previously reported as associated with lung function

The table below shows the results for the 279 variants that are reported as associated with lung function[20]. "Lung function trait" is the measure of lung function that the variant showed the strongest association with out of FEV_1, FVC, FEV_1/FVC and PEF. The beta and p value are the effect size of the variant and strength of association on the trait in the "Lung function trait" column. Rows highlighted in green are significantly associated with IPF risk after a Bonferroni correction for 279 tests ($p < 1.79 \times 10^{-4}$).

Chr	Position	rsid	Locus	Effect allele	Effect allele frequency	IPF OR	IPF risk p	Lung function trait	Lung function beta	Lung function p
1	6678864	rs9661802	*PHF13*	C	34.2%	1.00 [0.94, 1.08]	0.891	FEV_1/FVC	−0.025 [−0.030, −0.020]	5.56×10^{-23}
1	17308254	rs9435733	*MFAP2*	T	47.3%	1.03 [0.97, 1.10]	0.369	FEV_1/FVC	0.039 [0.034, 0.044]	5.95×10^{-61}
1	22612690	rs12737805	*MIR4418*	A	77.9%	1.06 [0.98, 1.14]	0.159	FEV_1	0.020 [0.015, 0.025]	6.57×10^{-13}
1	26775367	rs9438626	*DHDDS*	G	78.6%	1.00 [0.93, 1.08]	0.963	FVC	−0.018 [−0.023, −0.013]	8.06×10^{-10}
1	26796922	rs12096239	*DHDDS*	G	74.3%	1.02 [0.95, 1.10]	0.573	FEV_1	0.019 [0.014, 0.024]	2.08×10^{-12}
1	39995074	rs755249	LOC101929516	T	23.6%	1.05 [0.98, 1.14]	0.179	FEV_1/FVC	−0.024 [−0.029, −0.019]	9.82×10^{-18}
1	51243374	rs1416685	*FAF1*	C	41.5%	1.05 [0.98, 1.12]	0.170	FEV_1/FVC	0.020 [0.015, 0.025]	5.62×10^{-17}
1	60966772	rs72673461	LOC101926964	T	95.3%	1.07 [0.92, 1.25]	0.376	FEV_1/FVC	0.051 [0.040, 0.062]	3.12×10^{-20}
1	78387270	rs9661687	*NEXN*	T	86.5%	1.03 [0.94, 1.14]	0.539	FEV_1/FVC	−0.027 [−0.034, −0.02]	6.11×10^{-15}
1	92077097	rs1192415	*TGFBR3*	G	18.7%	1.07 [0.99, 1.17]	0.096	FEV_1/FVC	−0.044 [−0.050, −0.038]	2.28×10^{-47}
1	92106637	rs10874851	*TGFBR3*	A	48.9%	1.08 [1.01, 1.15]	0.030	FEV_1/FVC	−0.014 [−0.019, −0.009]	5.07×10^{-9}
1	92381483	rs11165787	*TGFBR3*	G	31.2%	1.05 [0.98, 1.12]	0.204	FEV_1/FVC	−0.024 [−0.029, −0.019]	1.63×10^{-21}
1	111737398	rs9970286	*DENND2D*	A	32.2%	1.06 [0.99, 1.14]	0.089	FEV_1/FVC	0.024 [0.019, 0.029]	1.92×10^{-21}
1	118911295	rs35043843	*SPAG17*	G	23.9%	1.03 [0.95, 1.11]	0.473	FVC	0.024 [0.019, 0.029]	4.12×10^{-18}
1	150249101	rs11205354	C1orf54	A	44.3%	1.04 [0.98, 1.11]	0.211	PEF	0.017 [0.012, 0.022]	2.49×10^{-11}
1	150547747	rs878471	*MCL1*	A	57.7%	1.15 [1.08, 1.23]	2.94×10^{-5}	FVC	−0.028 [−0.033, −0.023]	1.50×10^{-31}
1	155137395	rs141942982	*KRTCAP2*	G	89.3%	1.15 [1.04, 1.28]	0.008	FEV_1/FVC	0.036 [0.028, 0.044]	9.57×10^{-21}
1	178719306	rs4651005	*RALGPS2*	T	31.0%	1.07 [1.00, 1.15]	0.064	FEV_1	0.018 [0.013, 0.023]	1.38×10^{-13}
1	186090370	rs2146098	*MIR548F1*	A	64.5%	1.01 [0.95, 1.08]	0.702	FVC	−0.018 [−0.023, −0.013]	2.03×10^{-13}
1	186113852	rs17531405	*MIR548F1*	C	18.1%	1.06 [0.97, 1.15]	0.192	FEV_1/FVC	0.028 [0.022, 0.034]	2.80×10^{-19}
1	198898157	rs10919604	*MIR181A1HG*	G	40.1%	1.08 [1.01, 1.15]	0.024	FEV_1/FVC	−0.020 [−0.025, −0.015]	4.48×10^{-16}
1	200069216	rs2816992	*NR5A2*	G	41.5%	1.01 [0.95, 1.08]	0.776	FVC	0.016 [0.011, 0.021]	7.33×10^{-12}

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

1	201884647	rs4309038	*LMOD1*	G	56.2%	1.01 [0.94, 1.08]	0.813	FEV_1/FVC	−0.015 [−0.020, −0.010]	2.13×10^{-10}
1	204426295	rs1008833	*PIK3C2B*	A	85.3%	1.07 [0.97, 1.17]	0.172	PEF	−0.032 [−0.039, −0.025]	1.50×10^{-19}
1	215120596	rs556648	*CENPF/KCNK2*	A	22.1%	1.02 [0.94, 1.10]	0.618	FVC	0.015 [0.009, 0.021]	3.12×10^{-7}
1	218521609	rs2799098	*TGFB2*	A	83.1%	1.03 [0.94, 1.12]	0.512	FEV_1/FVC	−0.028 [−0.034, −0.022]	5.00×10^{-20}
1	218631452	rs6604614	*TGFB2*	C	71.1%	1.04 [0.97, 1.12]	0.262	PEF	−0.016 [−0.021, −0.011]	2.13×10^{-8}
1	218855029	rs28613267	*MIR548F3/TGFB2*	C	50.1%	1.05 [0.99, 1.12]	0.126	FEV_1	0.017 [0.012, 0.022]	9.19×10^{-13}
1	219483218	rs75128958	*LYPLAL1*	G	92.2%	1.12 [0.99, 1.26]	0.081	FEV_1/FVC	0.044 [0.035, 0.053]	2.33×10^{-23}
1	219853742	rs1338227	*RNU5F-1*	G	43.2%	1.07 [1.00, 1.14]	0.052	FEV_1/FVC	−0.024 [−0.029, −0.019]	3.92×10^{-24}
1	221204299	rs17009288	*HLX*	C	29.2%	1.01 [0.94, 1.08]	0.815	FVC	0.025 [0.020, 0.030]	3.68×10^{-22}
1	221631938	rs12757436	C1orf140/*DUSP10*	A	33.3%	1.00 [0.93, 1.07]	0.952	FVC	0.016 [0.011, 0.021]	1.76×10^{-10}
1	239857524	rs2355237	*CHRM3*	A	50.8%	1.02 [0.95, 1.09]	0.602	PEF	0.029 [0.024, 0.034]	9.42×10^{-31}
2	15906854	rs2544536	LOC101926966	C	52.3%	1.05 [0.99, 1.12]	0.122	FEV_1/FVC	0.024 [0.019, 0.029]	4.15×10^{-24}
2	18287623	rs55884799	*KCNS3*	T	81.5%	1.01 [0.93, 1.10]	0.754	FEV_1/FVC	−0.042 [−0.048, −0.036]	4.02×10^{-40}
2	18570024	rs6751968	*RDH14*	C	81.7%	1.00 [0.92, 1.09]	0.982	FVC	−0.025 [−0.031, −0.019]	1.27×10^{-16}
2	18702313	rs13430465	*RDH14*	T	7.7%	1.09 [0.96, 1.23]	0.175	FVC	0.037 [0.029, 0.045]	1.87×10^{-17}
2	24018480	rs13009582	*ATAD2B*	G	54.0%	1.01 [0.95, 1.08]	0.720	FVC	−0.016 [−0.021, −0.011]	1.52×10^{-11}
2	26842146	rs732990	*CIB4*	C	44.3%	1.02 [0.95, 1.09]	0.629	FVC	−0.016 [−0.021, −0.011]	3.72×10^{-11}
2	42243850	rs4952564	*PKDCC*	A	67.9%	1.02 [0.96, 1.10]	0.500	FVC	−0.017 [−0.022, −0.012]	6.97×10^{-12}
2	56096892	rs3791679	*EFEMP1*	G	22.9%	1.02 [0.94, 1.10]	0.598	FVC	−0.034 [−0.039, −0.029]	7.19×10^{-35}
2	102926362	rs12470864	*IL1RL1*	G	61.6%	1.07 [1.00, 1.15]	0.040	FEV_1/FVC	0.020 [0.015, 0.025]	1.04×10^{-16}
2	135672187	rs62168891	*CCNT2-AS1*	C	56.7%	1.03 [0.96, 1.10]	0.442	FVC	−0.019 [−0.024, −0.014]	1.08×10^{-14}
2	145797829	rs1406225	*TEX41*	G	71.6%	1.11 [1.03, 1.19]	0.007	FEV_1/FVC	0.020 [0.015, 0.025]	8.73×10^{-14}
2	157016257	rs72902177	LOC101929378	C	87.2%	1.06 [0.96, 1.16]	0.279	FEV_1	0.034 [0.027, 0.041]	1.76×10^{-22}
2	161276378	rs7424771	*RBMS1*	G	55.8%	1.03 [0.96, 1.10]	0.438	FEV_1	0.017 [0.012, 0.022]	6.57×10^{-13}
2	179260382	rs2304340	*MIR548N*	G	58.7%	1.02 [0.95, 1.09]	0.610	FEV_1	0.014 [0.009, 0.019]	3.72×10^{-9}
2	187530520	rs2084448	*ITGAV*	C	29.9%	1.04 [0.97, 1.11]	0.316	FEV_1/FVC	−0.020 [−0.025, −0.015]	4.65×10^{-14}
2	199723365	rs1249096	*SATB2*	G	43.2%	1.01 [0.95, 1.08]	0.744	FVC	−0.021 [−0.026, −0.016]	3.94×10^{-18}
2	201208692	rs985256	*SPATS2L*	A	21.9%	1.02 [0.95, 1.11]	0.571	FEV_1/FVC	0.018 [0.012, 0.024]	5.86×10^{-10}
2	202970250	rs12997625	*KIAA2012*	C	47.8%	1.09 [1.02, 1.16]	0.014	FVC	0.017 [0.012, 0.022]	8.88×10^{-13}

2	217614730	rs6435952	*IGFBP5*	T	85.6%	1.01 [0.92, 1.11]	0.875	FEV_1/FVC	−0.026 [−0.032, −0.020]	1.06×10^{-14}
2	218604356	rs4294980	*DIRC3*	G	21.0%	1.05 [0.97, 1.14]	0.254	FEV_1	−0.018 [−0.024, −0.012]	4.02×10^{-10}
2	218683154	rs2571445	*TNS1*	G	61.3%	1.07 [1.00, 1.14]	0.050	FEV_1	0.028 [0.023, 0.033]	7.24×10^{-33}
2	220382700	rs4674407	*ASIC4*	T	49.8%	1.00 [0.94, 1.07]	0.929	FVC	−0.015 [−0.020, −0.010]	2.84×10^{-10}
2	229502197	rs62201738	*PID1*	C	7.7%	1.05 [0.93, 1.19]	0.412	FEV_1/FVC	0.074 [0.065, 0.083]	9.45×10^{-63}
2	239441308	rs6710301	*TRAF3IP1*	A	14.4%	1.01 [0.92, 1.11]	0.779	FEV_1	0.024 [0.018, 0.030]	8.51×10^{-13}
2	239604970	rs6431620	LINC01107	T	78.1%	1.02 [0.95, 1.11]	0.567	FVC	0.018 [0.012, 0.024]	1.40×10^{-10}
2	239881309	rs4308141	*FLJ43879*	G	20.0%	1.08 [0.99, 1.17]	0.077	FEV_1/FVC	0.048 [0.042, 0.054]	3.58×10^{-59}
2	241837000	rs9973765 [a]	C2orf54	T	50.8%	1.01 [0.94, 1.08]	0.807	FVC	0.019 [0.014, 0.023]	1.19×10^{-13}
2	242495953	rs6733504	*BOK-AS1*	G	45.6%	1.02 [0.95, 1.09]	0.615	FVC	−0.019 [−0.024, −0.014]	7.39×10^{-16}
3	13787641	rs2974389	LINC00620	A	42.8%	1.07 [1.00, 1.14]	0.045	FEV_1	0.016 [0.011, 0.021]	1.70×10^{-12}
3	25179533	rs73048404	*RARB*	T	86.0%	1.03 [0.94, 1.14]	0.491	FVC	0.021 [0.015, 0.027]	2.10×10^{-10}
3	25520582	rs1529672	*RARB*	C	82.5%	1.04 [0.95, 1.13]	0.384	FEV_1/FVC	−0.042 [−0.048, −0.036]	1.73×10^{-41}
3	29469675	rs17666332	*RBMS3*	G	27.8%	1.01 [0.94, 1.08]	0.847	FEV_1/FVC	−0.027 [−0.032, −0.022]	9.22×10^{-24}
3	55152319	rs12715478	*CACNA2D3*	A	59.6%	1.06 [0.99, 1.13]	0.092	FEV_1/FVC	0.025 [0.020, 0.030]	1.35×10^{-24}
3	57879611	rs6445932	*SLMAP*	G	25.3%	1.03 [0.96, 1.11]	0.391	FEV_1	0.029 [0.024, 0.034]	3.82×10^{-26}
3	67455803	rs4132748	*SUCLG2*	C	70.4%	1.03 [0.96, 1.11]	0.387	FEV_1	0.020 [0.015, 0.025]	1.18×10^{-15}
3	71583177	rs35480566	*FOXP1*	A	57.0%	1.08 [1.01, 1.15]	0.028	FVC	−0.022 [−0.027, −0.017]	1.25×10^{-20}
3	73862616	rs586936	*PDZRN3-AS1*	A	40.1%	1.00 [0.94, 1.07]	0.907	FEV_1/FVC	−0.018 [−0.023, −0.013]	2.11×10^{-13}
3	98822050	rs12497779	*DCBLD2*	T	23.6%	1.04 [0.96, 1.12]	0.373	FVC	−0.032 [−0.037, −0.027]	1.85×10^{-30}
3	99420192	rs1610265	*MIR548G*	T	7.1%	1.08 [0.95, 1.23]	0.249	FVC	−0.038 [−0.047, −0.029]	8.81×10^{-18}
3	127931340	rs2999090	*EEFSEC*	A	87.6%	1.01 [0.91, 1.11]	0.905	FEV_1/FVC	−0.043 [−0.050, −0.036]	6.76×10^{-32}
3	158226886	rs12634907	*RSRC1*	A	66.2%	1.04 [0.97, 1.11]	0.267	FVC	0.026 [0.021, 0.031]	2.83×10^{-26}
3	165548529	rs1799807	*BCHE*	C	2.0%	1.05 [0.82, 1.33]	0.718	FEV_1/FVC	−0.06 [−0.077, −0.043]	8.59×10^{-12}
3	168709843	rs879394	LOC100507661	G	77.4%	1.06 [0.98, 1.15]	0.138	FEV_1	0.029 [0.024, 0.034]	1.36×10^{-25}
3	169295436	rs78101726	*MECOM*	A	85.5%	1.00 [0.91, 1.10]	0.945	FEV_1	0.033 [0.027, 0.039]	7.72×10^{-25}
3	185530290	rs6769511 [a]	*IGF2BP2*	C	31.4%	1.03 [0.96, 1.11]	0.390	FEV_1	−0.017 [−0.022, −0.013]	6.18×10^{-12}
4	7879027	rs62289340	*AFAP1*	T	43.7%	1.04 [0.97, 1.11]	0.295	FEV_1/FVC	0.017 [0.012, 0.022]	2.36×10^{-12}
4	56012149	rs12331869	*KDR*	G	82.0%	1.02 [0.94, 1.12]	0.578	FEV_1	0.018 [0.012, 0.024]	3.17×10^{-9}

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

4	75676529	rs62316310	*BTC*	G	73.5%	1.04 [0.97, 1.12]	0.285	FEV_1/FVC	−0.027 [−0.032, −0.022]	2.24×10^{-23}
4	79403952	rs11098196	*FRAS1*	G	48.3%	1.03 [0.96, 1.10]	0.398	FEV_1/FVC	0.020 [0.015, 0.025]	2.42×10^{-16}
4	89855495	rs2609279	*FAM13A*	T	22.6%	1.32 [1.22, 1.43]	1.93×10^{-12}	FEV_1/FVC	0.054 [0.048, 0.060]	2.08×10^{-76}
4	89869078	rs2869966	*FAM13A*	C	61.1%	1.25 [1.16, 1.33]	1.48×10^{-10}	FEV_1/FVC	0.042 [0.037, 0.047]	5.78×10^{-66}
4	106133184	rs6533183	*TET2*	C	35.9%	1.06 [0.99, 1.13]	0.095	FEV_1/FVC	0.030 [0.025, 0.035]	2.60×10^{-32}
4	106766430	rs11722225	*GSTCD*	C	7.0%	1.10 [0.97, 1.25]	0.139	FEV_1	0.073 [0.064, 0.082]	2.43×10^{-54}
4	106819053	rs34712979	*NPNT*	G	75.7%	1.00 [0.93, 1.08]	0.960	FEV_1/FVC	0.068 [0.063, 0.073]	4.18×10^{-134}
4	145330628	rs13109426	*HHIP-AS1*	G	40.2%	1.07 [1.00, 1.14]	0.060	FVC	0.023 [0.018, 0.028]	4.20×10^{-21}
4	145442364	rs13116999	*HHIP-AS1*	A	53.9%	1.12 [1.05, 1.19]	0.001	PEF	0.066 [0.061, 0.071]	2.54×10^{-153}
4	145506456	rs13141641	*HHIP-AS1*	C	40.2%	1.04 [0.97, 1.11]	0.294	FEV_1/FVC	0.070 [0.065, 0.075]	3.65×10^{-184}
4	145740898	rs2353940	*OTUD4/SMAD1*	T	75.1%	1.10 [1.02, 1.18]	0.017	PEF	0.038 [0.032, 0.044]	5.11×10^{-40}
5	609661	rs11739847	LOC100996325	G	80.9%	1.09 [1.00, 1.18]	0.046	FEV_1	0.021 [0.015, 0.027]	4.30×10^{-13}
5	33352738	rs268717	*TARS*	T	90.9%	1.06 [0.94, 1.18]	0.345	FVC	−0.034 [−0.042, −0.026]	3.23×10^{-17}
5	43976162	rs4866846	*NNT*	G	84.9%	1.01 [0.92, 1.11]	0.829	FEV_1	−0.028 [−0.034, −0.022]	2.47×10^{-17}
5	44367221	rs6859730	*FGF10*	T	67.1%	1.02 [0.96, 1.1]	0.494	FVC	−0.021 [−0.026, −0.016]	8.76×10^{-17}
5	52187038	rs12522114	*ITGA1*	A	26.4%	1.08 [1.00, 1.16]	0.053	FEV_1/FVC	−0.037 [−0.042, −0.032]	1.47×10^{-41}
5	53444498	rs2441026	*ARL15*	C	54.0%	1.04 [0.97, 1.11]	0.282	FVC	−0.018 [−0.023, −0.013]	7.79×10^{-14}
5	77396400	rs425102	*AP3B1*	T	76.0%	1.04 [0.96, 1.12]	0.338	FVC	0.021 [0.016, 0.026]	2.78×10^{-14}
5	95025146	rs987068	*SPATA9*	G	31.3%	1.13 [1.06, 1.22]	0.001	FEV_1/FVC	0.030 [0.025, 0.035]	1.46×10^{-30}
5	121410529	rs10059661	*LOX*	C	82.6%	1.07 [0.98, 1.16]	0.154	FEV_1/FVC	−0.031 [−0.037, −0.025]	1.78×10^{-22}
5	128767384	rs17163397	*ADAMTS19-AS1*	G	12.2%	1.09 [0.99, 1.21]	0.087	FEV_1/FVC	0.031 [0.024, 0.038]	3.30×10^{-17}
5	131421190	rs6898270 [a]	*P4HA2-AS1*	T	43.3%	1.07 [1.00, 1.15]	0.039	FVC	0.019 [0.014, 0.023]	7.02×10^{-15}
5	147856522	rs7733410	*HTR4*	G	56.0%	1.03 [0.96, 1.10]	0.376	FEV_1/FVC	−0.050 [−0.055, −0.045]	1.56×10^{-96}
5	148206885	rs1800888	*ADRB2*	C	98.8%	1.06 [0.78, 1.44]	0.711	FEV_1	0.084 [0.065, 0.103]	6.45×10^{-18}
5	148652302	rs11952673	*ABLIM3*	G	61.0%	1.02 [0.95, 1.09]	0.554	FEV_1	0.019 [0.014, 0.024]	1.35×10^{-14}
5	156908317	rs11134766	*CYFIP2*	C	93.5%	1.12 [0.98, 1.28]	0.094	FEV_1/FVC	0.063 [0.053, 0.073]	8.04×10^{-38}
5	156944199	rs11134789	*ADAM19*	C	65.4%	1.08 [1.01, 1.16]	0.032	FEV_1/FVC	0.041 [0.036, 0.046]	3.06×10^{-59}
5	170901463	rs10059996	*FGF18*	G	63.8%	1.09 [1.02, 1.17]	0.010	FEV_1/FVC	0.035 [0.030, 0.040]	1.53×10^{-42}
5	179598771	rs79898473	*RASGEF1C*	C	33.2%	1.00 [0.93, 1.07]	0.960	FEV_1/FVC	0.031 [0.026, 0.036]	2.31×10^{-33}

6	6741932	rs1294417	*LY86*	T	45.5%	1.00 [0.94, 1.07]	0.911	FEV_1/FVC	−0.031 [−0.036, −0.026]	3.93×10^{-39}
6	7563232	rs2076295	*DSP*	G	46.9%	1.46 [1.37, 1.56]	2.79×10^{-30}	FEV_1/FVC	0.023 [0.018, 0.028]	6.95×10^{-23}
6	7720059	rs12198986	*BMP6*	A	47.5%	1.02 [0.95, 1.08]	0.647	FVC	−0.023 [−0.028, −0.018]	2.17×10^{-22}
6	7797840	rs10498672	*BMP6*	G	17.2%	1.02 [0.94, 1.12]	0.616	FVC	−0.035 [−0.041, −0.029]	7.50×10^{-31}
6	22017543	rs13198081	*CASC15*	C	36.1%	1.12 [1.05, 1.20]	0.001	FEV_1/FVC	0.030 [0.025, 0.035]	3.07×10^{-33}
6	28301099	rs7752448	*ZNF184*	G	11.1%	1.16 [1.04, 1.28]	0.006	PEF	−0.055 [−0.062, −0.048]	1.21×10^{-48}
6	32151443	rs2070600	*AGER*	T	5.8%	1.07 [0.93, 1.24]	0.320	FEV_1/FVC	0.145 [0.135, 0.155]	3.00×10^{-189}
6	32612396	rs9273084 [a]	*HLA-DQB1*	T	40.1%	1.07 [0.97, 1.19]	0.185	FEV_1/FVC	−0.047 [−0.053, −0.041]	9.79×10^{-57}
6	34188892	rs9689096	*HMGA1*	C	5.9%	1.03 [0.90, 1.18]	0.683	FVC	0.036 [0.026, 0.046]	1.99×10^{-13}
6	44447598	rs9357446	*CDC5L*	A	52.1%	1.02 [0.95, 1.09]	0.577	FVC	−0.015 [−0.020, −0.010]	1.04×10^{-10}
6	45530471	rs12202314	*RUNX2*	C	32.5%	1.01 [0.94, 1.08]	0.867	FEV_1/FVC	0.021 [0.016, 0.026]	2.17×10^{-16}
6	45622748	rs9472541	*RUNX2*	T	71.0%	1.02 [0.95, 1.10]	0.527	FVC	0.015 [0.010, 0.020]	2.47×10^{-9}
6	56336406	rs2894837	*RNU6-71P*	G	37.0%	1.03 [0.96, 1.10]	0.357	FEV_1	−0.018 [−0.023, −0.013]	9.22×10^{-13}
6	73663814	rs13206405	*KCNQ5*	A	20.1%	1.09 [1.01, 1.19]	0.033	FEV_1/FVC	0.034 [0.028, 0.040]	4.67×10^{-31}
6	109268050	rs2798641	*ARMC2*	C	81.9%	1.06 [0.97, 1.16]	0.173	FEV_1/FVC	0.045 [0.039, 0.051]	3.89×10^{-48}
6	126990392	rs6918725	*MIR588*	G	52.1%	1.06 [0.99, 1.13]	0.100	FVC	0.021 [0.016, 0.026]	4.73×10^{-19}
6	134339265	rs2627237	*SLC2A12*	G	40.4%	1.02 [0.95, 1.09]	0.567	FEV_1	−0.014 [−0.019, −0.009]	3.50×10^{-9}
6	140271357	rs1102077	LOC100507477	A	76.4%	1.01 [0.93, 1.09]	0.872	FEV_1	0.022 [0.017, 0.027]	4.21×10^{-15}
6	142560957	rs9385988	*VTA1*	G	27.2%	1.09 [1.01, 1.17]	0.026	FEV_1	0.028 [0.023, 0.033]	1.40×10^{-26}
6	142688969	rs17280293	*GPR126*	A	97.0%	1.07 [0.88, 1.29]	0.511	FEV_1/FVC	−0.18 [−0.195, −0.165]	2.34×10^{-131}
6	142745883	rs7753012	*GPR126*	G	31.7%	1.01 [0.94, 1.09]	0.712	FEV_1/FVC	0.071 [0.066, 0.076]	4.71×10^{-165}
7	7256490	rs4318980	*C1GALT1*	G	59.0%	1.02 [0.96, 1.09]	0.510	FEV_1/FVC	0.017 [0.012, 0.022]	9.08×10^{-13}
7	15506007	rs4721442	*AGMO*	T	83.0%	1.03 [0.94, 1.12]	0.505	FVC	0.022 [0.016, 0.028]	7.21×10^{-12}
7	15872324	rs4721457	*MEOX2-AS1*	C	15.1%	1.03 [0.94, 1.13]	0.528	FEV_1/FVC	−0.024 [−0.03, −0.018]	1.73×10^{-13}
7	26848830	rs559233	*SKAP2*	T	48.7%	1.01 [0.95, 1.08]	0.727	FEV_1	0.017 [0.012, 0.022]	7.80×10^{-13}
7	27182329	rs62454414	*HOXA-AS3*	T	87.2%	1.01 [0.92, 1.12]	0.780	FVC	0.021 [0.014, 0.028]	1.31×10^{-9}
7	28200097	rs1513272	*JAZF1*	T	49.1%	1.03 [0.97, 1.10]	0.310	FEV_1	0.020 [0.015, 0.025]	1.11×10^{-17}
7	46448518	rs17232687	*IGFBP3*	C	50.4%	1.00 [0.94, 1.07]	0.942	FVC	0.018 [0.013, 0.023]	6.81×10^{-15}
7	84569510	rs12707691	*SEMA3D*	G	33.4%	1.02 [0.96, 1.10]	0.492	FEV_1	0.020 [0.015, 0.025]	1.67×10^{-16}

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

7	99692993	rs2261360	*ZKSCAN1*	T	25.3%	1.25 [1.16, 1.34]	5.50×10^{-9}	FEV_1/FVC	0.022 [0.017, 0.027]	8.74×10^{-15}
7	116431427	rs193686	*MET*	C	32.2%	1.01 [0.94, 1.08]	0.868	FEV_1/FVC	0.018 [0.013, 0.023]	4.07×10^{-12}
7	156127246	rs12698403	LOC285889	G	57.0%	1.09 [1.02, 1.16]	0.016	FEV_1	0.027 [0.022, 0.032]	6.42×10^{-31}
8	9018590	rs330939	*PPP1R3B*	T	60.9%	1.01 [0.94, 1.09]	0.742	FEV_1/FVC	0.023 [0.018, 0.028]	4.46×10^{-21}
8	11823332	rs4128298	*DEFB136*	C	26.9%	1.01 [0.93, 1.10]	0.755	FEV_1	0.017 [0.012, 0.022]	3.48×10^{-11}
8	70367248	rs7465401	LOC100505739	C	28.1%	1.03 [0.96, 1.11]	0.385	FEV_1	0.021 [0.016, 0.026]	9.09×10^{-16}
8	145504343	rs7838717	*BOP1*	T	36.5%	1.05 [0.98, 1.13]	0.133	FVC	−0.023 [−0.028, −0.018]	6.47×10^{-21}
9	1568941	rs771662	*DMRT2/SMARCA2*	C	65.9%	1.03 [0.96, 1.10]	0.450	FVC	0.016 [0.011, 0.021]	1.08×10^{-10}
9	4120648	rs1570203	*GLIS3*	A	52.8%	1.01 [0.95, 1.08]	0.699	FEV_1/FVC	0.025 [0.020, 0.030]	5.78×10^{-25}
9	18013733	rs7041139	*SH3GL2*	C	67.8%	1.07 [1.00, 1.15]	0.047	FEV_1	0.017 [0.012, 0.022]	3.09×10^{-12}
9	23587027	rs1107677	*FLJ35282/ELAVL2*	T	49.3%	1.00 [0.94, 1.07]	0.931	FEV_1/FVC	0.022 [0.017, 0.027]	3.88×10^{-20}
9	98266855	rs28446321	*PTCH1*	A	9.2%	1.02 [0.91, 1.15]	0.693	FEV_1/FVC	−0.052 [−0.06, −0.044]	4.72×10^{-36}
9	98878881	rs72743974	LOC158434	G	16.8%	1.02 [0.94, 1.12]	0.590	FEV_1/FVC	0.023 [0.017, 0.029]	3.98×10^{-13}
9	101632854	rs57649467	*GALNT12*	A	39.2%	1.08 [1.01, 1.15]	0.035	FEV_1/FVC	0.018 [0.013, 0.023]	5.39×10^{-13}
9	109483517	rs1491106	*TMEM38B/ZNF462*	G	62.7%	1.14 [1.06, 1.22]	2.20×10^{-4}	FEV_1/FVC	−0.024 [−0.029, −0.019]	2.81×10^{-23}
9	119234058	rs10983184	*ASTN2*	C	35.3%	1.02 [0.95, 1.09]	0.670	FEV_1/FVC	−0.027 [−0.032, −0.022]	9.05×10^{-28}
9	131943843	rs967497	*IER5L*	A	31.0%	1.02 [0.95, 1.09]	0.576	FEV_1	0.015 [0.010, 0.020]	2.79×10^{-9}
9	139100413	rs7024579	*QSOX2*	T	31.7%	1.00 [0.93, 1.07]	0.996	FVC	−0.023 [−0.028, −0.018]	4.48×10^{-20}
9	139259349	rs4073153	*DNLZ*	G	43.7%	1.02 [0.96, 1.09]	0.528	FVC	−0.014 [−0.019, −0.009]	8.64×10^{-9}
10	12278021	rs7090277	*CDC123*	A	53.2%	1.03 [0.96, 1.10]	0.383	FEV_1/FVC	0.041 [0.036, 0.046]	3.97×10^{-67}
10	30268770	rs7914842	*KIAA1462*	A	57.1%	1.01 [0.95, 1.08]	0.710	PEF	0.016 [0.011, 0.021]	1.02×10^{-10}
10	34480582	rs1274475	*PARD3*	A	38.8%	1.02 [0.95, 1.09]	0.630	FEV_1/FVC	0.017 [0.012, 0.022]	8.30×10^{-12}
10	64998971	rs7082066	*JMJD1C*	A	19.5%	1.02 [0.94, 1.11]	0.618	FEV_1	0.022 [0.016, 0.028]	2.20×10^{-13}
10	69962954	rs10998018	*MYPN*	A	49.2%	1.01 [0.94, 1.07]	0.838	FVC	−0.022 [−0.027, −0.017]	2.43×10^{-21}
10	75580014	rs7098573	*CAMK2G*	G	27.2%	1.09 [1.01, 1.17]	0.030	FEV_1	0.025 [0.020, 0.030]	2.75×10^{-21}
10	75639578	rs60820984	*CAMK2G*	C	81.2%	1.04 [0.96, 1.14]	0.329	PEF	0.020 [0.014, 0.026]	6.38×10^{-10}
10	77119039	rs1259605	*COMTD1/ZNF503-AS1*	C	24.6%	1.01 [0.94, 1.09]	0.746	FVC	0.012 [0.007, 0.017]	1.19×10^{-5}
10	78312002	rs2637254	C10orf11	A	51.6%	1.02 [0.96, 1.09]	0.465	FEV_1	−0.028 [−0.033, −0.023]	3.91×10^{-34}
10	81706324	rs721917	*SFTPD*	A	57.7%	1.00 [0.94, 1.07]	0.998	FEV_1/FVC	0.019 [0.014, 0.024]	1.65×10^{-15}

10	105639611	rs11191841	*OBFC1*	C	50.8%	1.15 [1.08, 1.23]	2.05×10^{-5}	FEV_1	0.017 [0.012, 0.022]	6.21×10^{-13}
10	124297637	rs4279944	*HTRA1*	T	15.1%	1.09 [0.99, 1.19]	0.079	FEV_1/FVC	0.022 [0.015, 0.029]	2.18×10^{-10}
11	35308988	rs10836366	*SLC1A2*	C	24.9%	1.04 [0.97, 1.12]	0.288	FEV_1/FVC	−0.019 [−0.024, −0.014]	2.33×10^{-12}
11	43690717	rs17596617	*HSD17B12*	T	32.0%	1.02 [0.95, 1.10]	0.501	FVC	−0.020 [−0.025, −0.015]	3.58×10^{-15}
11	45244903	rs10838435	*PRDM11*	C	15.6%	1.06 [0.97, 1.16]	0.219	FEV_1	0.021 [0.015, 0.027]	1.46×10^{-10}
11	62370155	rs71490394	*EML3*	A	37.0%	1.01 [0.94, 1.08]	0.760	FEV_1	0.026 [0.021, 0.031]	1.66×10^{-27}
11	73036179	rs2027761	*ARHGEF17*	C	88.6%	1.00 [0.91, 1.11]	0.942	FEV_1/FVC	−0.037 [−0.044, −0.030]	1.31×10^{-22}
11	86448839	rs11234768	*PRSS23*	T	85.0%	1.03 [0.94, 1.12]	0.588	FEV_1/FVC	0.030 [0.024, 0.036]	5.07×10^{-20}
11	126009500	rs541601	*RPUSD4*	T	19.0%	1.04 [0.96, 1.13]	0.356	FEV_1/FVC	−0.024 [−0.030, −0.018]	5.29×10^{-15}
12	2908330	rs56196860	*FKBP4*	A	2.8%	1.09 [0.89, 1.33]	0.421	FVC	−0.053 [−0.067, −0.039]	1.45×10^{-14}
12	4243749	rs12811814	*CCND2-AS1*	T	45.9%	1.03 [0.97, 1.10]	0.337	FEV_1	0.015 [0.010, 0.020]	2.57×10^{-10}
12	19808912	rs10841302	*AEBP2*	G	44.9%	1.05 [0.98, 1.12]	0.177	FEV_1/FVC	−0.017 [−0.022, −0.012]	2.29×10^{-12}
12	28588242	rs7977418	*CCDC91*	T	54.4%	1.03 [0.96, 1.10]	0.381	FVC	0.038 [0.033, 0.043]	9.57×10^{-59}
12	56396768	rs1689510	*RAB5B*	G	66.8%	1.02 [0.95, 1.09]	0.559	FEV_1	0.015 [0.010, 0.020]	5.57×10^{-10}
12	57527283	rs11172113	*LRP1*	T	59.9%	1.04 [0.98, 1.12]	0.202	FEV_1/FVC	−0.023 [−0.028, −0.018]	7.04×10^{-21}
12	65075332	rs1244869	*RASSF3*	G	36.5%	1.01 [0.94, 1.08]	0.810	FEV_1/FVC	−0.015 [−0.020, −0.010]	6.16×10^{-10}
12	65793153	rs12825748	*MSRB3*	G	69.1%	1.01 [0.94, 1.08]	0.794	FEV_1	−0.020 [−0.025, −0.015]	6.27×10^{-15}
12	66409367	rs11176001	*MIR6074*	A	13.1%	1.02 [0.92, 1.12]	0.754	FEV_1	−0.029 [−0.036, −0.022]	4.88×10^{-17}
12	85719906	rs56390486	*ALX1/RASSF9*	A	28.7%	1.01 [0.94, 1.09]	0.756	PEF	0.020 [0.015, 0.025]	1.72×10^{-12}
12	94194890	rs9788269	*CRADD*	A	74.4%	1.03 [0.96, 1.12]	0.389	FVC	−0.014 [−0.019, −0.009]	3.97×10^{-7}
12	95554771	rs113745635	*FGD6*	T	21.5%	1.01 [0.93, 1.10]	0.759	FEV_1/FVC	−0.028 [−0.034, −0.022]	2.36×10^{-21}
12	96242109	rs7970544	*SNRPF*	G	80.9%	1.10 [1.01, 1.20]	0.022	FEV_1/FVC	−0.044 [−0.050, −0.038]	1.45×10^{-46}
12	102824921	rs972936	*IGF1*	C	73.7%	1.06 [0.99, 1.15]	0.100	PEF	−0.029 [−0.035, −0.023]	1.85×10^{-23}
12	114669870	rs2701110	*TBX5*	C	83.1%	1.00 [0.92, 1.09]	0.953	FEV_1	−0.026 [−0.032, −0.020]	1.91×10^{-16}
12	115201436	rs10850377	*TBX3*	A	34.7%	1.06 [0.99, 1.14]	0.097	FEV_1	0.020 [0.015, 0.025]	4.02×10^{-15}
12	115501127	rs35505	*TBX3*	A	68.5%	1.01 [0.94, 1.09]	0.727	FVC	0.022 [0.017, 0.027]	1.92×10^{-18}
13	44820608	rs9533803	*MIR8079*	C	78.7%	1.03 [0.95, 1.12]	0.431	FEV_1/FVC	0.026 [0.020, 0.032]	2.92×10^{-19}
13	50707087	rs2812208	*DLEU1*	C	2.1%	1.25 [1.00, 1.57]	0.052	FEV_1	0.061 [0.045, 0.077]	4.95×10^{-14}
13	71647588	rs803765	LINC00348	C	64.6%	1.08 [1.01, 1.15]	0.029	FVC	0.024 [0.019, 0.029]	1.74×10^{-23}

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

13	80467235	rs4885681	LINC00382	C	27.4%	1.00 [0.93, 1.08]	0.898	FEV_1	−0.019 [−0.024, −0.014]	1.83×10^{-12}
13	99665512	rs11620380	*DOCK9*	C	89.3%	1.04 [0.94, 1.16]	0.466	FEV_1/FVC	0.027 [0.019, 0.035]	4.61×10^{-12}
13	109918493	rs9634470	*MYO16*	T	74.0%	1.10 [1.02, 1.18]	0.013	FEV_1/FVC	−0.021 [−0.026, −0.016]	2.73×10^{-14}
14	23429729	rs1951121	*HAUS4*	G	39.6%	1.03 [0.97, 1.10]	0.340	FEV_1/FVC	−0.019 [−0.024, −0.014]	7.55×10^{-15}
14	54346010	rs74053129	*MIR5580*	A	9.9%	1.07 [0.96, 1.20]	0.210	FEV_1/FVC	0.039 [0.031, 0.047]	2.15×10^{-22}
14	54419106	rs35107139	*BMP4*	C	42.4%	1.00 [0.94, 1.07]	0.929	FEV_1/FVC	−0.032 [−0.037, −0.027]	3.40×10^{-36}
14	74817418	rs10141786	*VRTN*	A	40.2%	1.05 [0.99, 1.13]	0.123	FVC	0.021 [0.016, 0.026]	9.53×10^{-19}
14	84338431	rs1756281	LINC00911	G	30.9%	1.01 [0.94, 1.08]	0.853	FEV_1/FVC	−0.024 [−0.029, −0.019]	1.38×10^{-19}
14	92512143	rs11160037	*TRIP11*	G	37.9%	1.10 [1.02, 1.17]	0.008	FEV_1	0.018 [0.013, 0.023]	4.69×10^{-13}
14	93098339	rs11621587	*RIN3*	G	81.8%	1.01 [0.93, 1.10]	0.826	FVC	−0.036 [−0.042, −0.030]	1.39×10^{-32}
15	40397191	rs34245505	*BMF*	G	19.8%	1.05 [0.97, 1.14]	0.234	FVC	−0.021 [−0.027, −0.015]	1.56×10^{-12}
15	40716253	rs2304645	*IVD*	C	53.0%	1.31 [1.22, 1.39]	7.34×10^{-16}	FEV_1	−0.015 [−0.020, −0.010]	2.93×10^{-11}
15	41255396	rs4924525	*CHAC1*	A	52.6%	1.05 [0.99, 1.13]	0.111	FVC	−0.017 [−0.022, −0.012]	3.45×10^{-13}
15	41840238	rs2012453	*RPAP1*	G	58.6%	1.05 [0.98, 1.12]	0.153	FEV_1/FVC	−0.024 [−0.029, −0.019]	4.26×10^{-23}
15	41953211	rs56383987	*MGA*	C	95.0%	1.15 [0.99, 1.34]	0.066	FEV_1/FVC	0.036 [0.026, 0.046]	7.08×10^{-12}
15	49409527	rs79234094	*COPS2*	G	74.2%	1.10 [1.02, 1.18]	0.016	FEV_1/FVC	−0.027 [−0.032, −0.022]	3.18×10^{-23}
15	49706145	rs35251997	*FAM227B*	T	7.2%	1.11 [0.98, 1.26]	0.102	FEV_1/FVC	0.051 [0.042, 0.060]	2.82×10^{-27}
15	63866877	rs62012772	*USP3*	C	18.1%	1.03 [0.95, 1.12]	0.485	FEV_1/FVC	0.029 [0.023, 0.035]	2.42×10^{-20}
15	67491274	rs12917612	*AAGAB*	C	76.7%	1.01 [0.94, 1.09]	0.795	FVC	0.023 [0.018, 0.028]	2.04×10^{-16}
15	71612514	rs1441358	*THSD4*	T	66.5%	1.07 [1.00, 1.15]	0.045	FEV_1/FVC	0.064 [0.059, 0.069]	4.12×10^{-145}
15	71803450	rs62015883	*THSD4*	C	83.1%	1.05 [0.96, 1.14]	0.300	FEV_1	0.021 [0.015, 0.027]	1.16×10^{-11}
15	73833600	rs7176074	*REC114*	G	95.2%	1.00 [0.86, 1.17]	0.990	FEV_1/FVC	−0.034 [−0.045, −0.023]	6.60×10^{-10}
15	84274591	rs1896797	*SH3GL3*	A	49.8%	1.02 [0.96, 1.09]	0.510	FEV_1/FVC	0.029 [0.024, 0.034]	2.48×10^{-34}
16	3583173	rs3751837	*CLUAP1*	T	21.4%	1.02 [0.94, 1.11]	0.626	FVC	−0.031 [−0.036, −0.026]	9.07×10^{-28}
16	4361138	rs56104880	*GLIS2-AS1*	C	30.7%	1.06 [0.98, 1.13]	0.136	FEV_1/FVC	−0.020 [−0.025, −0.015]	5.29×10^{-15}
16	10136889	rs11074547	*GRIN2A*	T	73.4%	1.02 [0.95, 1.10]	0.576	FVC	−0.017 [−0.022, −0.012]	1.99×10^{-10}
16	10740982	rs78442819	*TEKT5*	C	19.7%	1.04 [0.96, 1.13]	0.302	FEV_1/FVC	−0.036 [−0.042, −0.030]	2.25×10^{-31}
16	28870962	rs12446589	*IL27*	A	39.4%	1.01 [0.94, 1.08]	0.812	FEV_1	−0.013 [−0.018, −0.008]	2.76×10^{-8}
16	50188929	rs76219171	*PAPD5*	A	6.0%	1.08 [0.95, 1.24]	0.248	FVC	−0.035 [−0.045, −0.025]	2.09×10^{-12}

16	53935407	rs35420030	*FTO*	T	95.0%	1.04 [0.90, 1.21]	0.587	FEV_1/FVC	−0.045 [−0.055, −0.035]	3.08×10^{-17}
16	58063513	rs11648508	*MMP15*	G	30.2%	1.06 [0.98, 1.14]	0.124	FEV_1/FVC	−0.033 [−0.038, −0.028]	9.86×10^{-39}
16	69891510	rs8047194	*WWP2*	T	49.1%	1.02 [0.95, 1.09]	0.627	FEV_1	−0.021 [−0.026, −0.016]	6.70×10^{-19}
16	75411445	rs11858992	*CFDP1*	A	40.9%	1.01 [0.95, 1.08]	0.736	FEV_1/FVC	0.038 [0.033, 0.043]	4.83×10^{-55}
16	78225633	rs2345443	*WWOX*	G	67.9%	1.02 [0.95, 1.09]	0.657	FEV_1	−0.022 [−0.027, −0.017]	3.03×10^{-18}
16	86403821	rs12918140	LINC00917	C	11.0%	1.14 [1.02, 1.26]	0.016	FEV_1/FVC	−0.027 [−0.034, −0.020]	6.72×10^{-13}
16	86579223	rs6539952	*MTHFSD*	C	74.6%	1.08 [1.00, 1.16]	0.052	FEV_1	0.017 [0.012, 0.022]	3.50×10^{-10}
17	3882613	rs8082036	*ATP2A3*	G	48.5%	1.03 [0.96, 1.10]	0.413	FEV_1/FVC	−0.024 [−0.029, −0.019]	7.31×10^{-24}
17	6469793	rs4796334	*PITPNM3*	G	50.5%	1.02 [0.95, 1.09]	0.600	FEV_1	0.014 [0.009, 0.019]	7.43×10^{-9}
17	7163350	rs1215	*CLDN7*	A	84.7%	1.00 [0.91, 1.11]	0.936	FVC	0.022 [0.016, 0.028]	9.64×10^{-11}
17	7448457	rs4968200	*TNFSF12-TNFSF13*	G	85.8%	1.02 [0.93, 1.12]	0.686	FEV_1	0.022 [0.016, 0.028]	4.54×10^{-11}
17	15959714	rs9652828 [a]	*NCOR1*	T	53.9%	1.03 [0.97, 1.10]	0.320	FVC	−0.015 [−0.019, −0.010]	3.43×10^{-10}
17	28072327	rs2244592	*SSH2*	A	46.9%	1.01 [0.94, 1.08]	0.790	FEV_1/FVC	−0.032 [−0.037, −0.027]	4.60×10^{-42}
17	29210595	rs62070648	*SUZ12P1*	G	73.2%	1.01 [0.94, 1.09]	0.839	FVC	−0.021 [−0.026, −0.016]	5.28×10^{-15}
17	36915540	rs35246838	*PSMB3*	T	86.8%	1.02 [0.92, 1.12]	0.728	FEV_1/FVC	0.039 [0.032, 0.046]	1.41×10^{-27}
17	37504933	rs8069451	*FBXL20*	C	26.0%	1.03 [0.96, 1.11]	0.442	FVC	−0.020 [−0.025, −0.015]	7.29×10^{-14}
17	43940021	rs79412431	*MAPT-AS1*	G	79.0%	1.36 [1.26, 1.48]	3.83×10^{-14}	FEV_1	0.043 [0.037, 0.049]	3.10×10^{-49}
17	46552229	rs12945803	*LOC101927166*	C	21.8%	1.03 [0.96, 1.12]	0.403	FVC	−0.020 [−0.025, −0.015]	1.88×10^{-12}
17	54195453	rs28519449	*ANKFN1*	T	40.1%	1.00 [0.94, 1.07]	0.974	FVC	0.021 [0.016, 0.026]	1.27×10^{-18}
17	59286644	rs8068952	*BCAS3*	G	22.7%	1.03 [0.96, 1.12]	0.420	FEV_1/FVC	0.028 [0.022, 0.034]	1.21×10^{-22}
17	62497964	rs77672322	*DDX5*	C	98.0%	1.15 [0.89, 1.49]	0.277	FVC	0.045 [0.030, 0.060]	3.02×10^{-9}
17	62686730	rs11653958	*SMURF2*	G	26.0%	1.00 [0.93, 1.08]	0.979	FEV_1/FVC	−0.020 [−0.025, −0.015]	9.17×10^{-13}
17	68976415	rs6501431	*CASC17*	C	21.7%	1.01 [0.93, 1.09]	0.831	FVC	−0.017 [−0.023, −0.011]	1.12×10^{-9}
17	69201811	rs6501455	*CASC17*	A	51.0%	1.05 [0.98, 1.12]	0.178	FEV_1	0.030 [0.025, 0.035]	1.28×10^{-36}
17	69371318	rs996865	*CASC17*	C	92.4%	1.06 [0.94, 1.20]	0.320	FEV_1/FVC	0.048 [0.039, 0.057]	1.82×10^{-25}
17	73525670	rs9892893	*TSEN54*	T	26.6%	1.07 [0.99, 1.15]	0.073	FEV_1	0.020 [0.015, 0.025]	2.08×10^{-13}
17	79952944	rs59606152	*ASPSCR1*	C	89.0%	1.10 [0.98, 1.23]	0.113	FVC	−0.037 [−0.045, −0.029]	9.16×10^{-21}
18	8801351	rs513953	*MTCL1*	A	25.0%	1.01 [0.93, 1.08]	0.892	FEV_1	−0.027 [−0.032, −0.022]	1.24×10^{-24}
18	10078071	rs8089099	*VAPA*	G	72.7%	1.01 [0.94, 1.09]	0.703	FEV_1/FVC	−0.024 [−0.029, −0.019]	1.52×10^{-18}

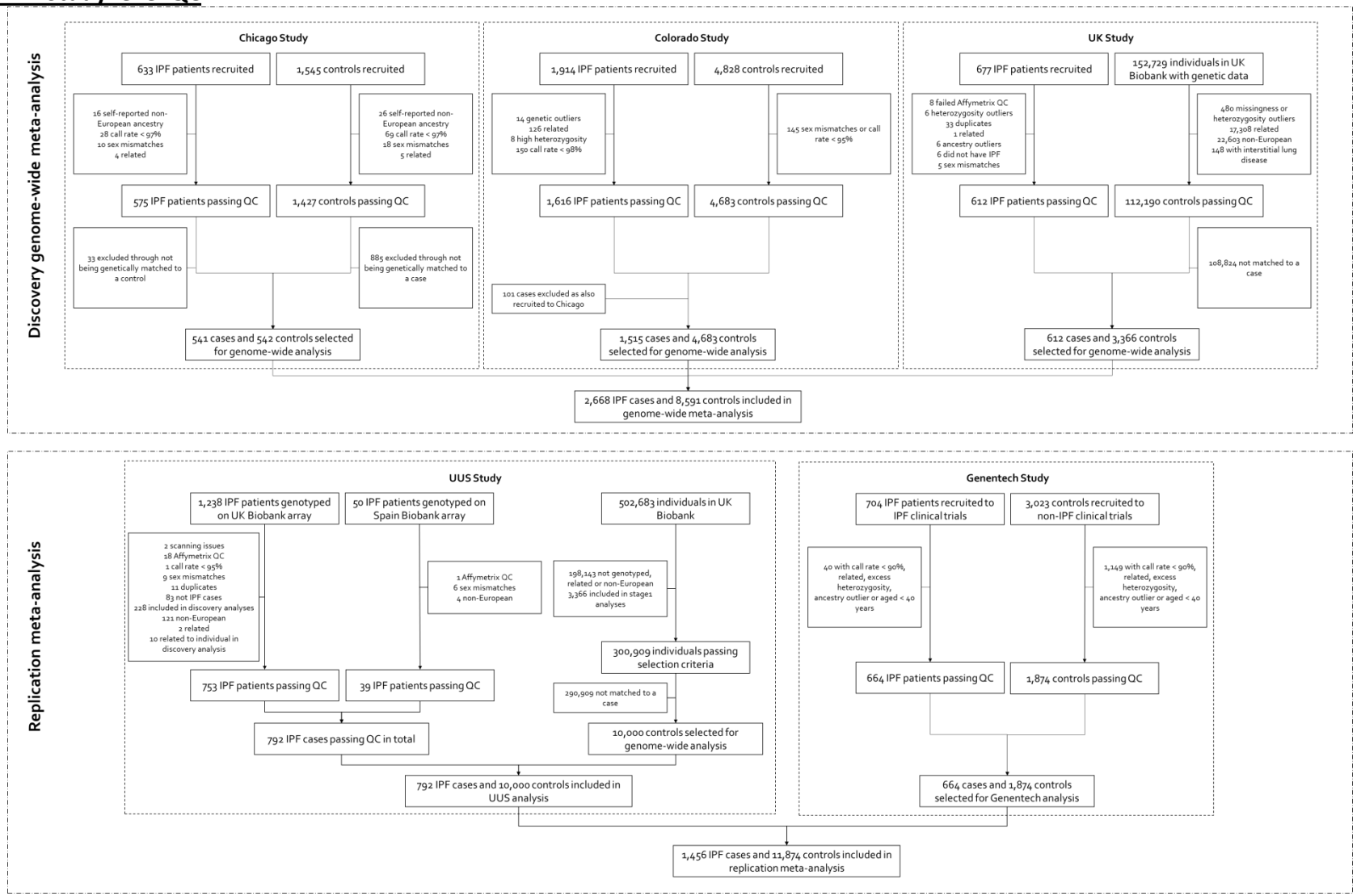
AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

18	19816712	rs1985511	*GATA6*	T	54.8%	1.00 [0.94, 1.07]	0.973	FEV_1/FVC	−0.016 [−0.021, −0.011]	5.60×10^{-12}
18	20234336	rs11082051	*CTAGE1/RBBP8*	A	51.8%	1.06 [1.00, 1.14]	0.069	FEV_1	0.013 [0.008, 0.018]	3.66×10^{-8}
18	20708321	rs9947743	*CABLES1*	A	78.9%	1.04 [0.96, 1.13]	0.306	FEV_1	−0.020 [−0.025, −0.015]	1.24×10^{-12}
18	21074255	rs303752	C18orf8	A	41.3%	1.01 [0.94, 1.08]	0.774	FVC	−0.016 [−0.021, −0.011]	6.97×10^{-12}
18	22290711	rs1668091	LOC729950	C	31.2%	1.02 [0.95, 1.09]	0.617	FVC	0.017 [0.012, 0.022]	5.24×10^{-12}
18	42827898	rs9807668	*SLC14A2*	T	9.7%	1.02 [0.92, 1.14]	0.667	FEV_1	0.029 [0.021, 0.037]	1.39×10^{-13}
18	51022606	rs12607758	*DCC*	C	40.5%	1.03 [0.96, 1.10]	0.446	FVC	−0.013 [−0.018, −0.008]	1.78×10^{-8}
18	53566471	rs2202572	LOC101927273	C	67.1%	1.01 [0.94, 1.08]	0.873	FVC	−0.016 [−0.021, −0.011]	7.01×10^{-11}
19	10819967	rs11085744	*QTRT1*	C	43.8%	1.06 [0.99, 1.13]	0.074	FEV_1	0.015 [0.010, 0.020]	7.83×10^{-11}
19	31829613	rs9636166	*TSHZ3*	C	13.0%	1.07 [0.97, 1.18]	0.194	FEV_1/FVC	−0.036 [−0.043, −0.029]	3.66×10^{-23}
19	36881643	rs2967516	*ZFP82*	G	28.8%	1.02 [0.95, 1.10]	0.518	FVC	0.015 [0.010, 0.020]	3.72×10^{-9}
19	41117300	rs34093919	*LTBP4*	A	1.4%	1.02 [0.77, 1.37]	0.869	FEV_1/FVC	0.154 [0.133, 0.175]	1.69×10^{-47}
20	6626218	rs2145272	*BMP2*	A	63.4%	1.01 [0.94, 1.08]	0.826	FVC	0.026 [0.021, 0.031]	1.42×10^{-26}
20	10745545	rs6032942	LOC101929395	C	23.1%	1.03 [0.95, 1.11]	0.498	FEV_1	0.017 [0.012, 0.022]	3.47×10^{-10}
20	25282608	rs2236180	*ABHD12*	C	18.9%	1.04 [0.96, 1.13]	0.350	FEV_1	−0.021 [−0.027, −0.015]	1.02×10^{-12}
20	30858967	rs4413223	C20orf112	G	82.8%	1.04 [0.95, 1.13]	0.426	FEV_1/FVC	0.023 [0.017, 0.029]	1.30×10^{-13}
20	34025756	rs143384	*UQCC1*	A	58.5%	1.05 [0.98, 1.12]	0.159	FVC	0.024 [0.019, 0.029]	1.03×10^{-23}
20	45486817	rs12481092	*EYA2*	T	26.8%	1.03 [0.96, 1.11]	0.451	FVC	0.026 [0.021, 0.031]	1.56×10^{-22}
20	62372706	rs4809221	*SLC2A4RG*	A	69.0%	1.15 [1.07, 1.23]	1.82×10^{-4}	FVC	−0.029 [−0.034, −0.024]	5.81×10^{-31}
21	35368402	rs12627254	LINC00649	G	86.9%	1.05 [0.96, 1.16]	0.286	FEV_1/FVC	−0.036 [−0.043, −0.029]	6.85×10^{-24}
21	35675966	rs62213732	*KCNE2*	T	63.4%	1.06 [0.99, 1.13]	0.118	FEV_1/FVC	0.025 [0.020, 0.030]	9.34×10^{-24}
22	18448113	rs1978968	*MICAL3*	C	76.4%	1.03 [0.96, 1.12]	0.396	FEV_1	−0.029 [−0.034, −0.024]	9.09×10^{-27}
22	20790723	rs9610955	*SCARF2*	C	19.6%	1.05 [0.97, 1.14]	0.247	FEV_1	−0.019 [−0.025, −0.013]	6.94×10^{-11}
22	28181399	rs2283847	*MN1*	T	55.5%	1.01 [0.95, 1.08]	0.684	FEV_1/FVC	−0.022 [−0.027, −0.017]	3.62×10^{-19}
22	50867711	rs113111175	*PPP6R2*	C	87.8%	1.01 [0.92, 1.12]	0.806	FEV_1	−0.022 [−0.029, −0.015]	1.11×10^{-9}

[a] Where the top lung function variant was not included in the IPF discovery GWAS, results are presented for proxy variants.

Supplementary Figures

Figure E1 - Study level QC



AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Figure E2 – Number of overlapping variants between studies included in the discovery genome-wide meta-analysis

Venn diagram showing the number of variants in each study and the amount of overlapping variants between studies used for the discovery genome-wide meta-analysis.

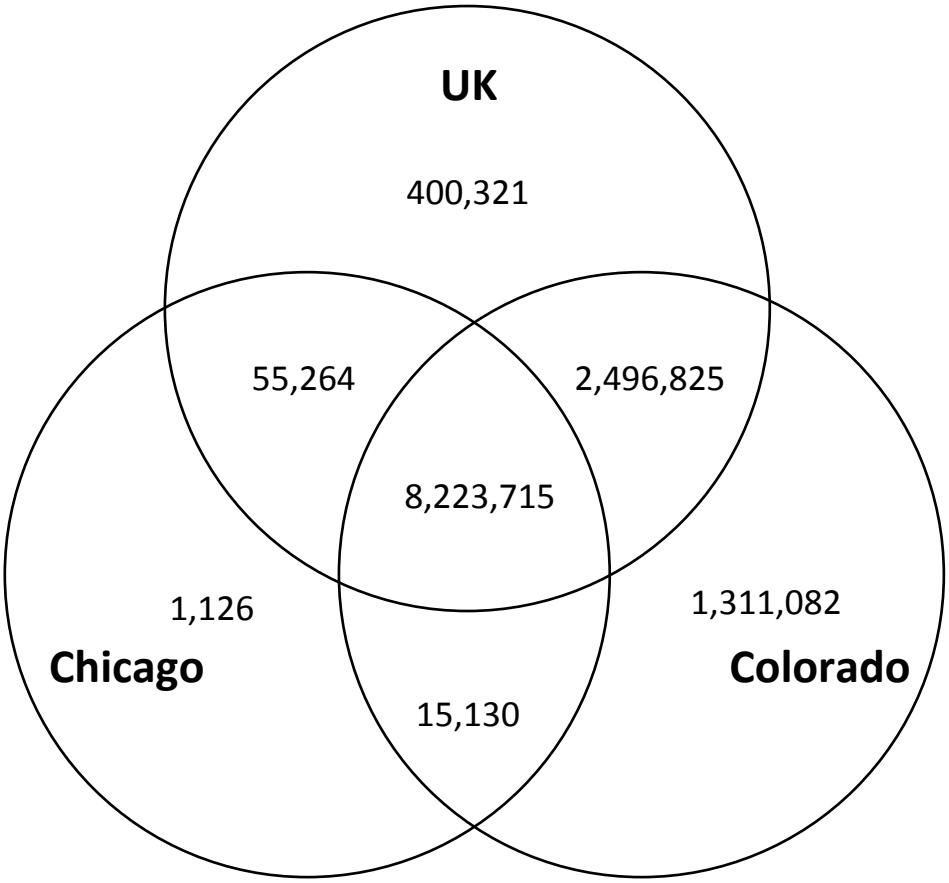
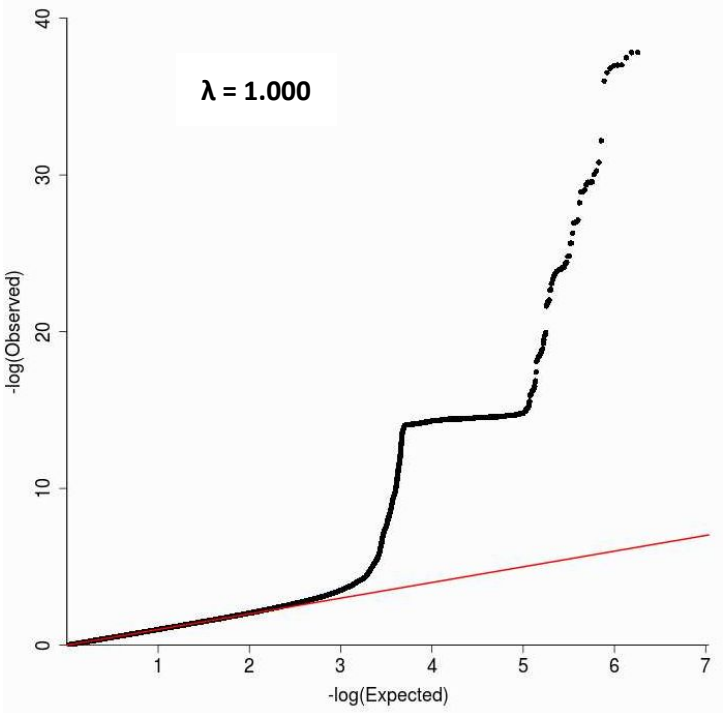
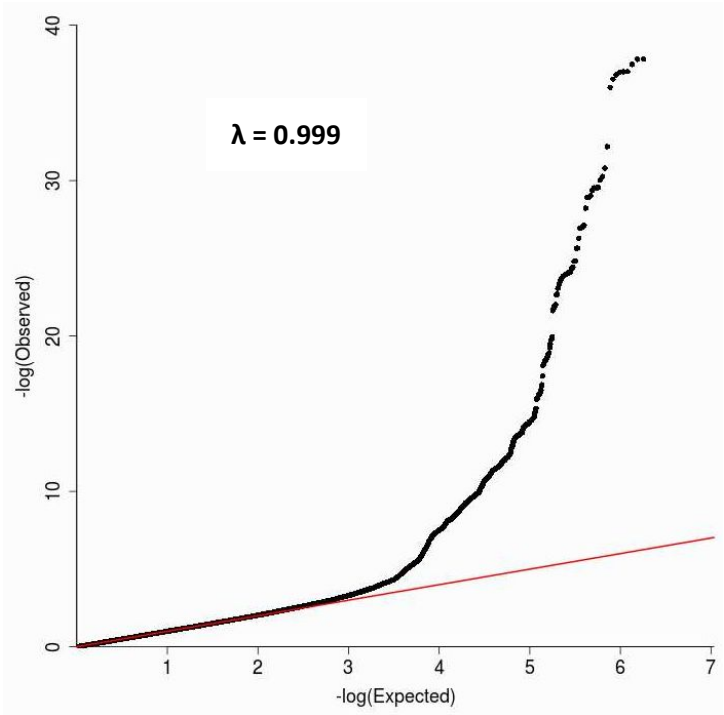


Figure E3 - QQ plot for discovery genome-wide meta-analysis

QQ plot for the genome-wide analysis with expected $-\log_{10}(P \text{ value})$ on the x axis and observed $-\log(P \text{ value})$ on the y axis. The red line shows where the expected equals the observed. To aid in viewing the figure, the y axis has been truncated at $-\log_{10}(P \text{ value}) = 40$. Variant rs35705950 (found in the *MUC5B* promoter region) has $-\log_{10}(P \text{ value}) = 202.9$.



The QQ plot above shows an increasing line before plateuing with a large number of variants with the same p value. This is due to the inversion region around *KANSL1* and *MAPT* where a large number of variants in very high LD show similar strengths of association with IPF. Below is the QQ plot excluding this inversion region on chromosome 17. To aid in viewing the figure, the y axis has been truncated at 40.

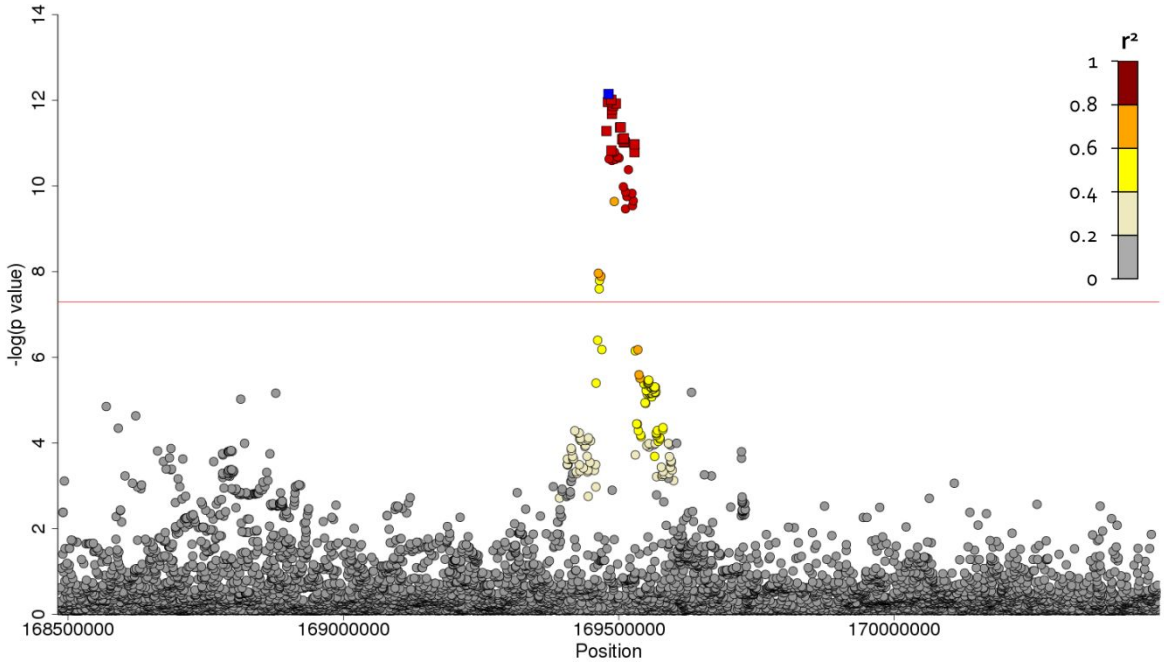


AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

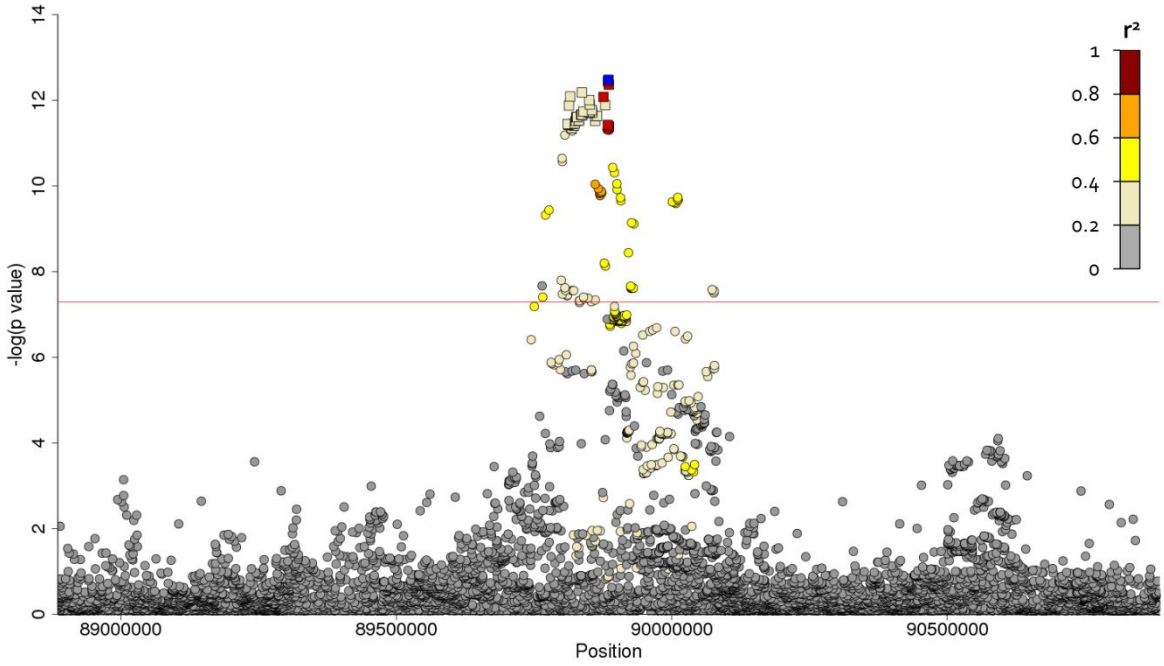
Figure E4 - Region plots for all 17 previously reported association signals

Region plots for each of the 17 previously reported signals. The x axis shows chromosomal position and y axis the −log(*P* value). The sentinel (or previously reported variant if there is no signal in the meta-analysis) is shown in blue with all other variants coloured by LD with the sentinel variant. Credible sets were calculated for each signal and variants in the credible set are shown by squares. The red horizontal line shows $P = 5 \times 10^{-8}$.

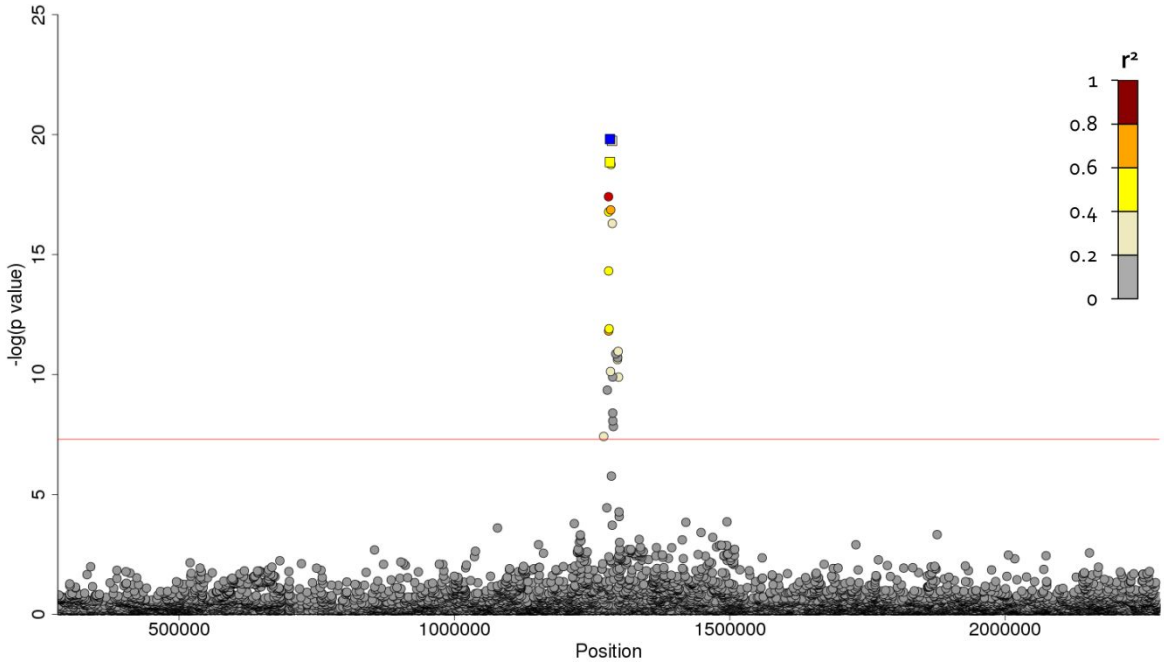
i) *TERC*



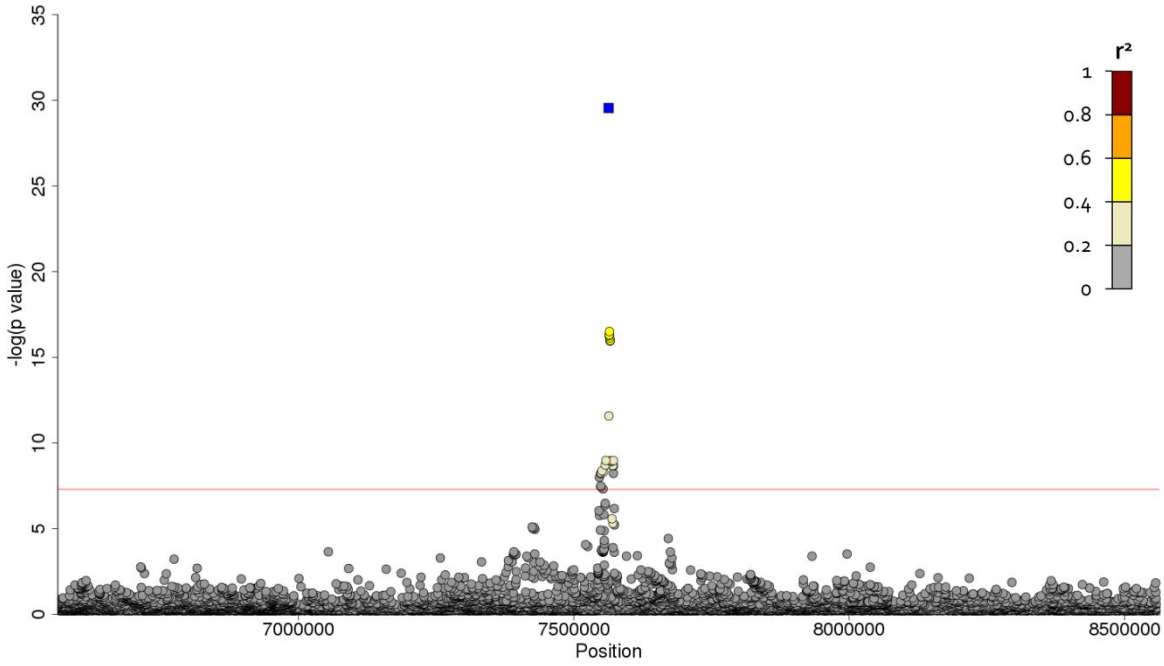
ii) *FAM13A*



iii) *TERT*

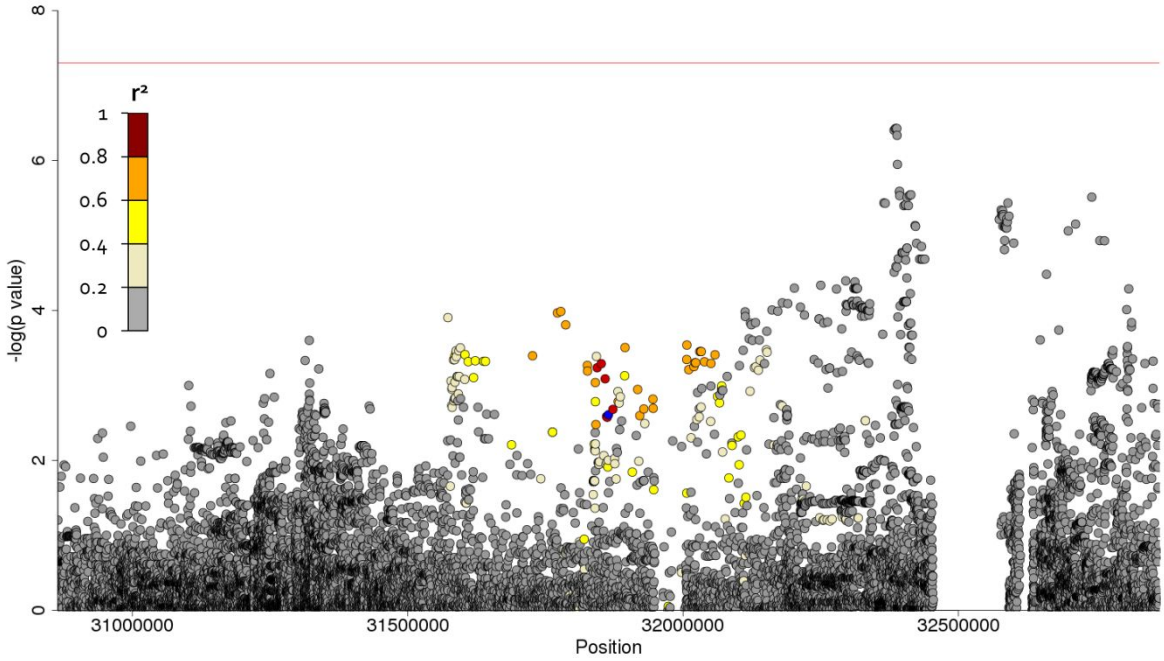


iv) *DSP*

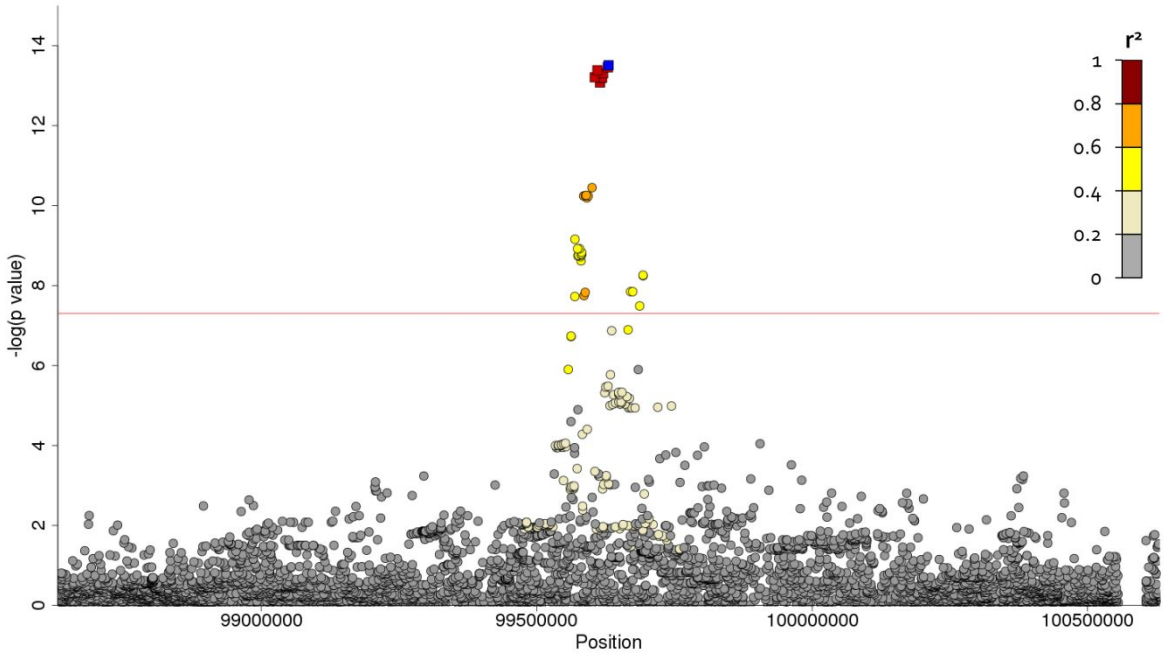


AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

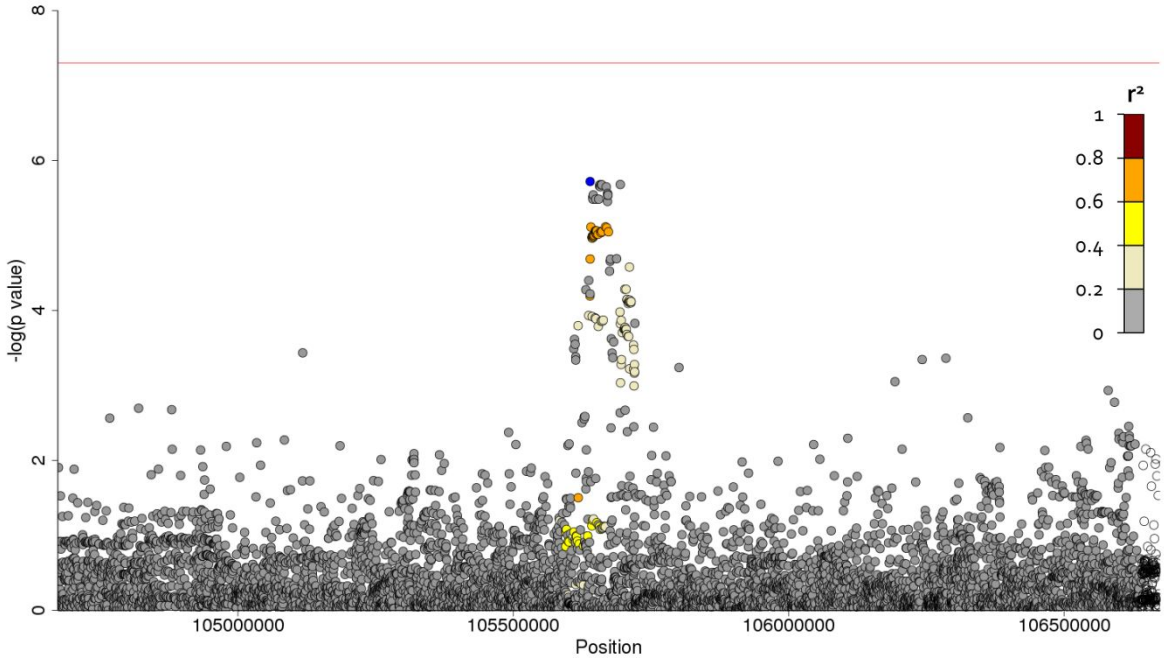
v) *EHMT2*



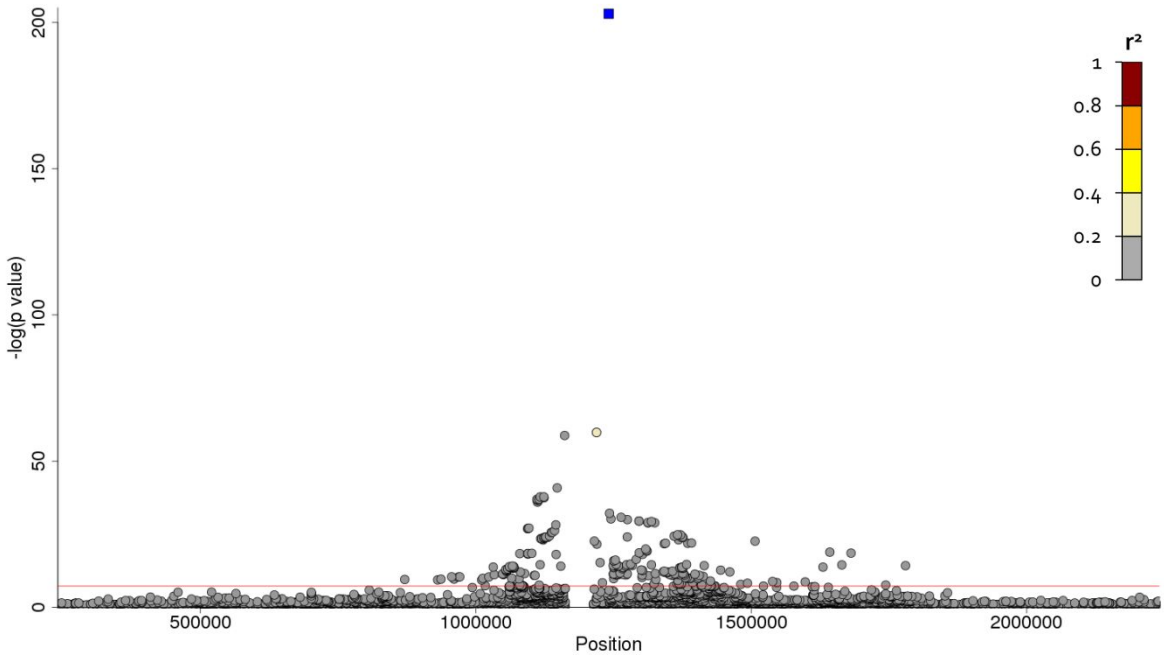
vi) 7q22.1



vii) *OBFC1*

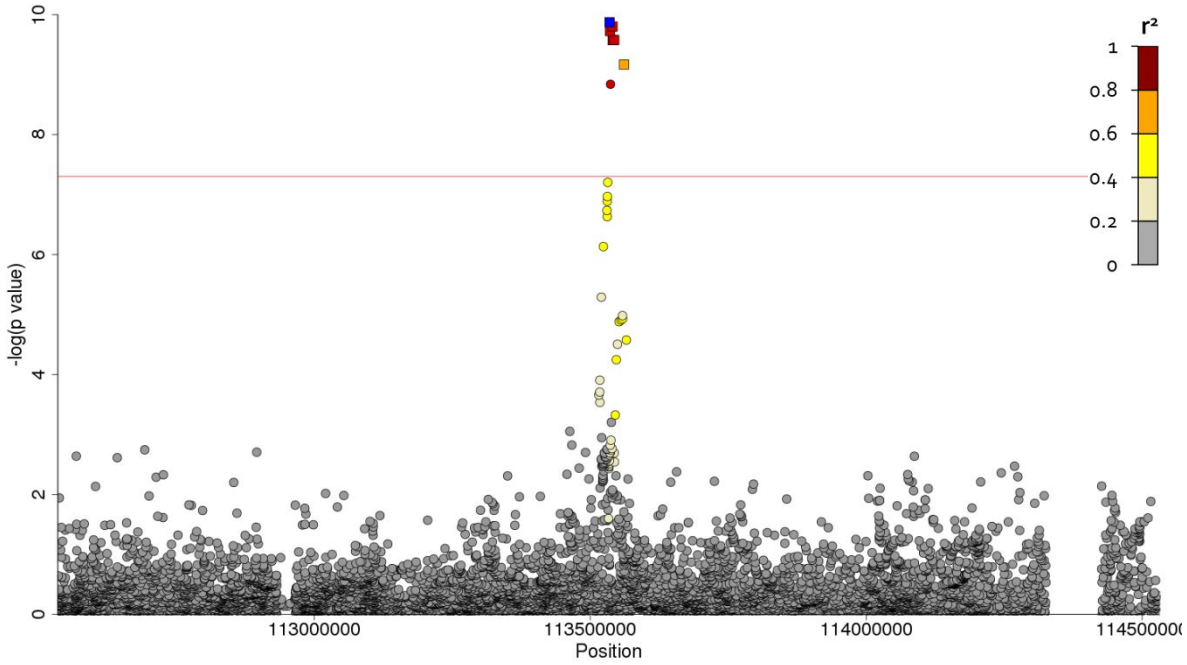


viii) *MUC5B*

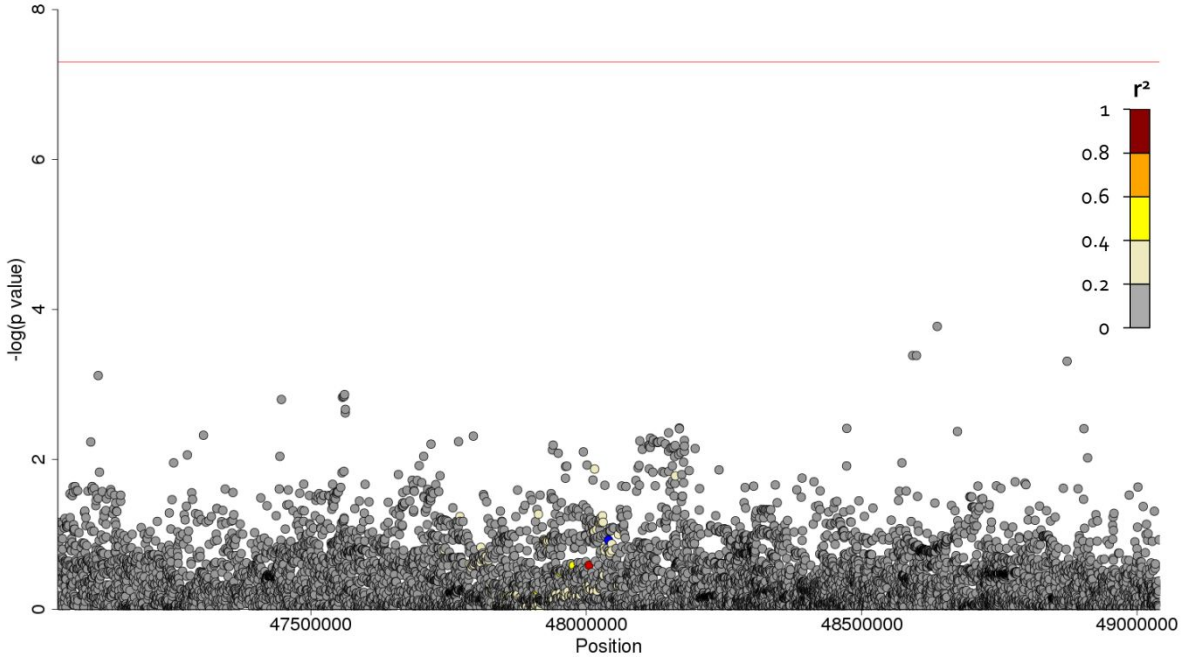


AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

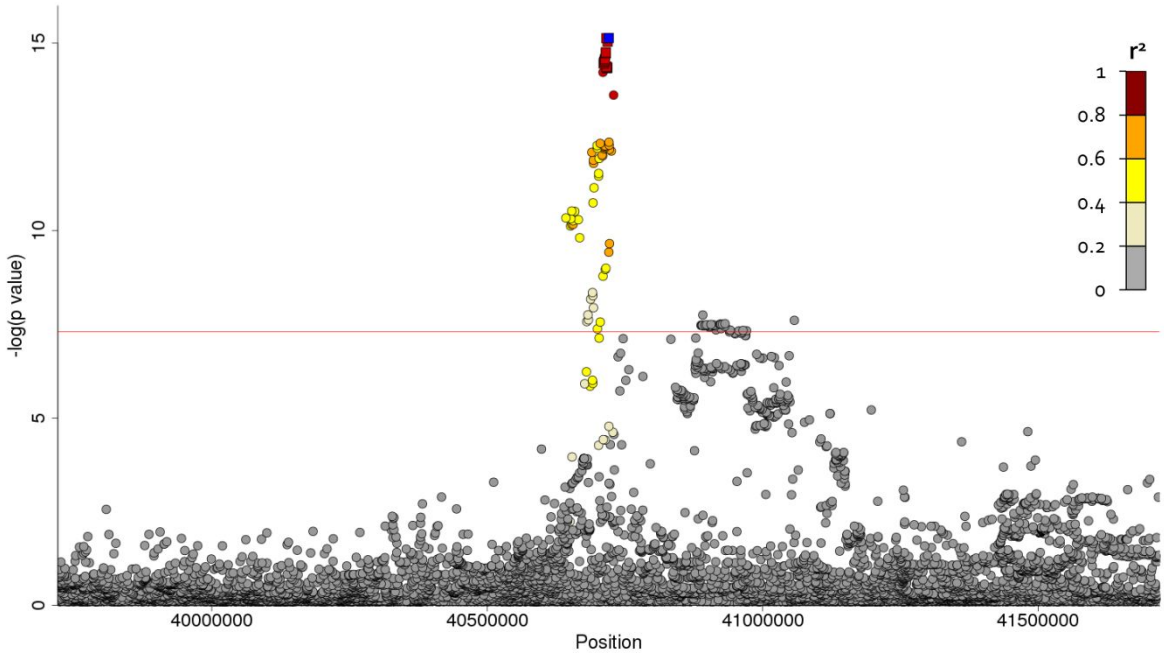
ix) *ATP11A*



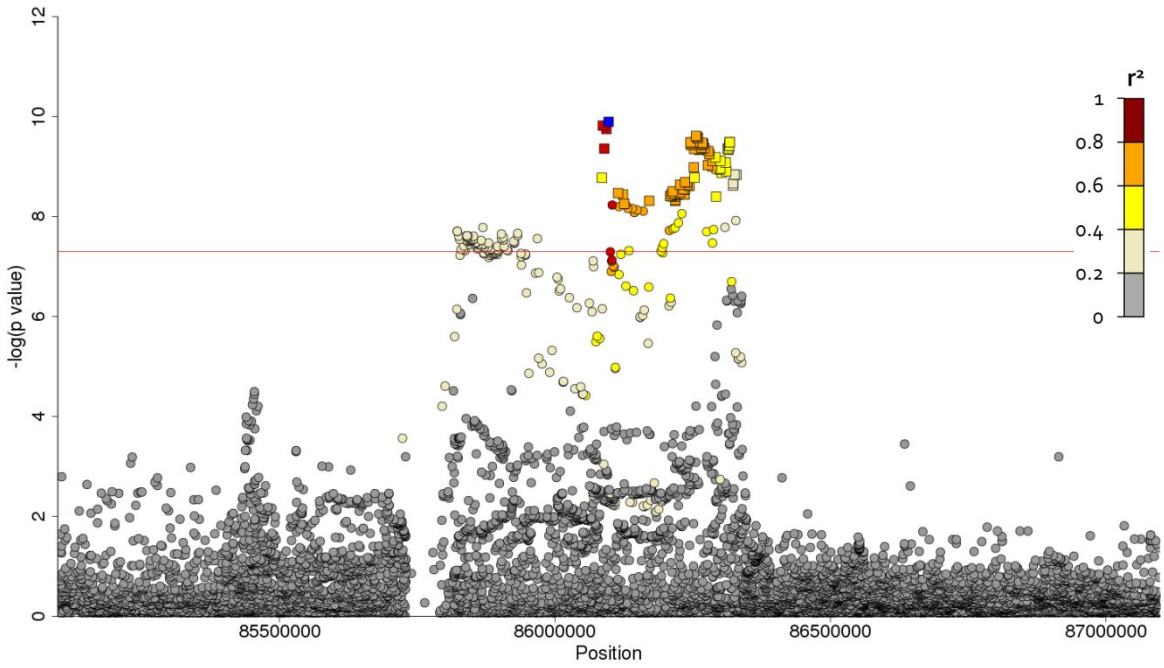
x) *MDGA2*



xi) *IVD*

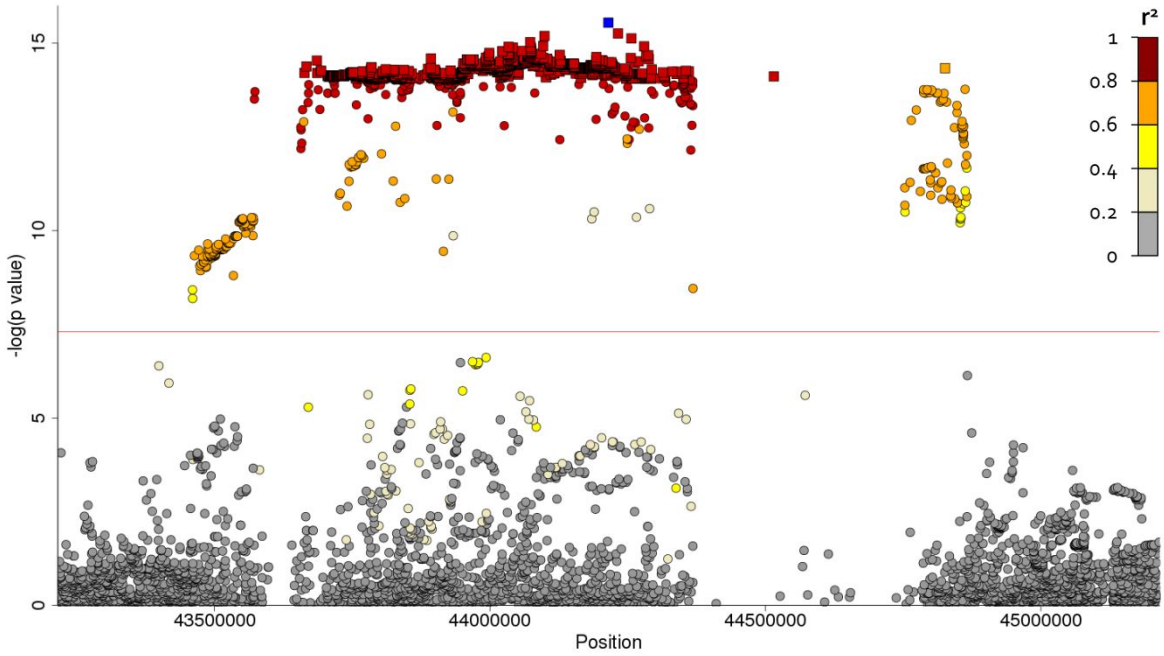


xii) *AKAP13*



AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

xiii) *MAPT*



xiv) *DPP9*

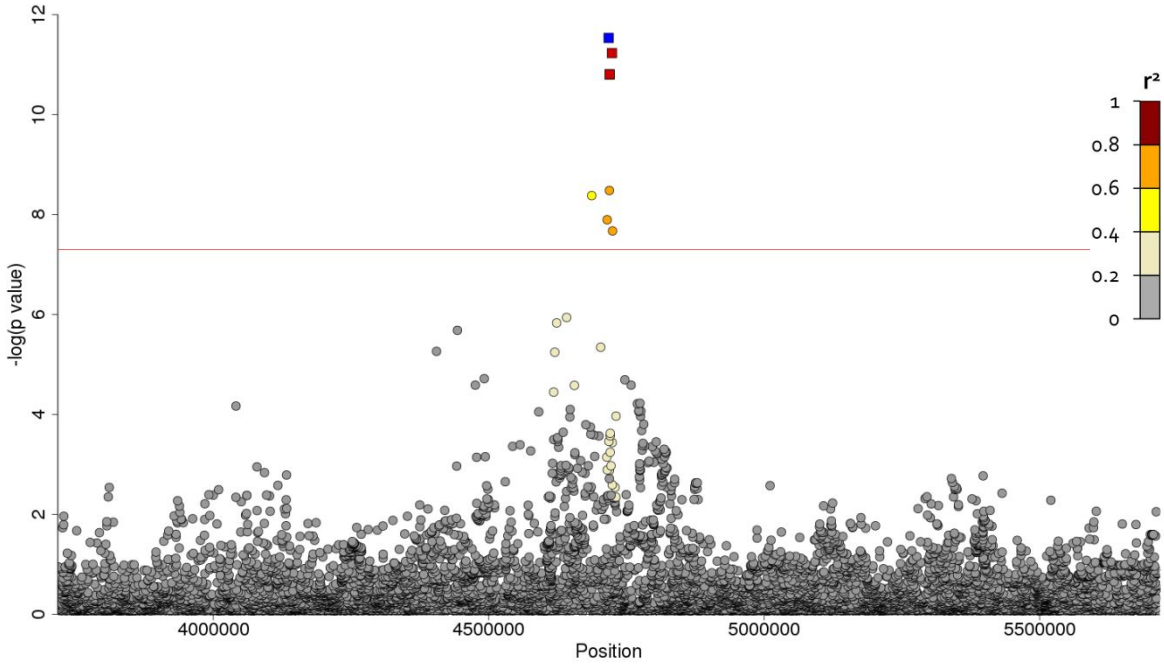
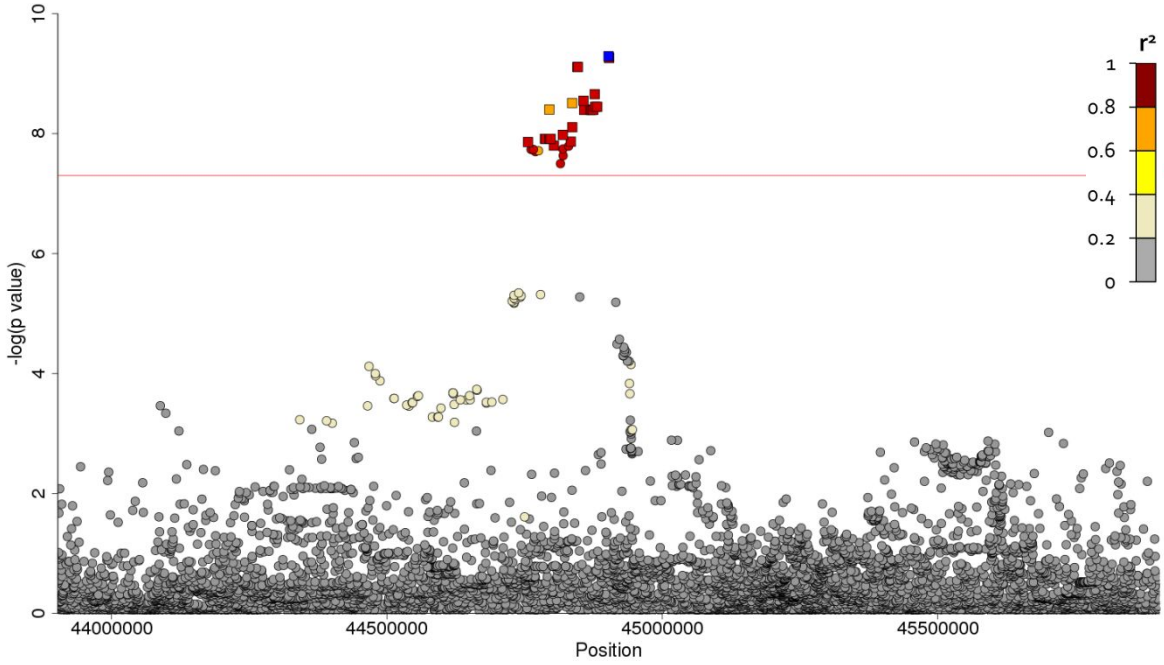


Figure E5 - Region plots and conditional analyses for the five novel IPF association signals in the discovery genome-wide analysis

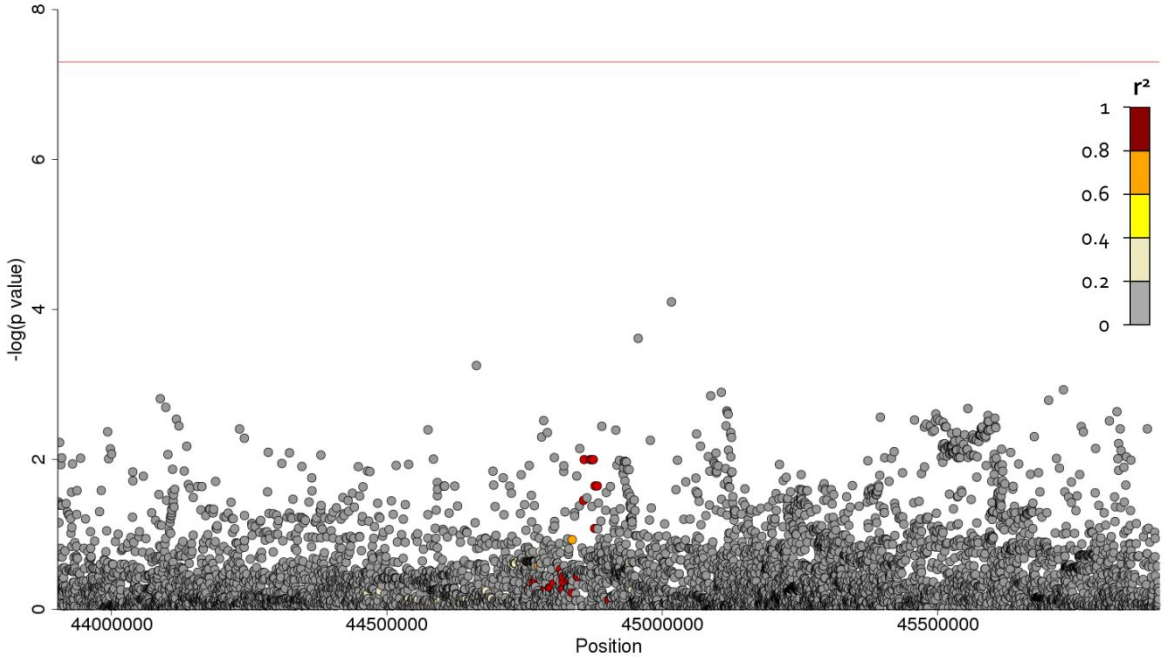
Region plots for each of the five novel signals in discovery analysis and after conditioning on the sentinel variant. The x axis shows chromosomal position and y axis the −log(*P* value). The sentinel is shown in blue with all other variants coloured by LD with the sentinel variant. Credible sets were calculated and variants in the credible set are shown by squares. The red horizontal line shows $P = 5 \times 10^{-8}$.

i) Chromosome 3

Region plot



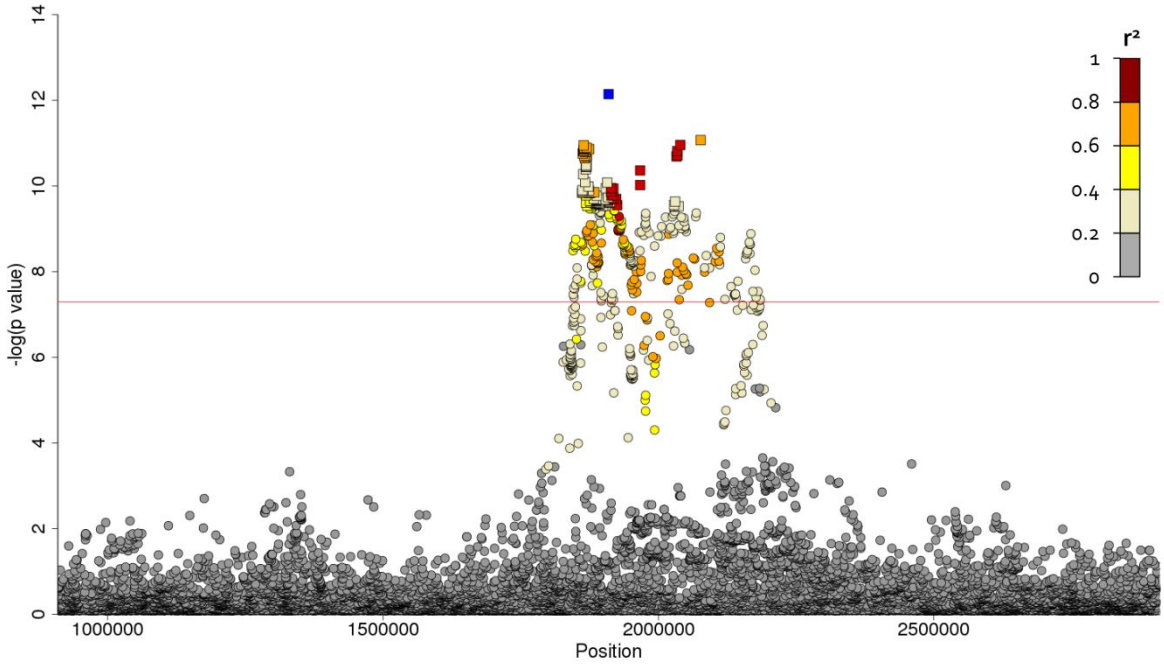
After conditioning on sentinel variant



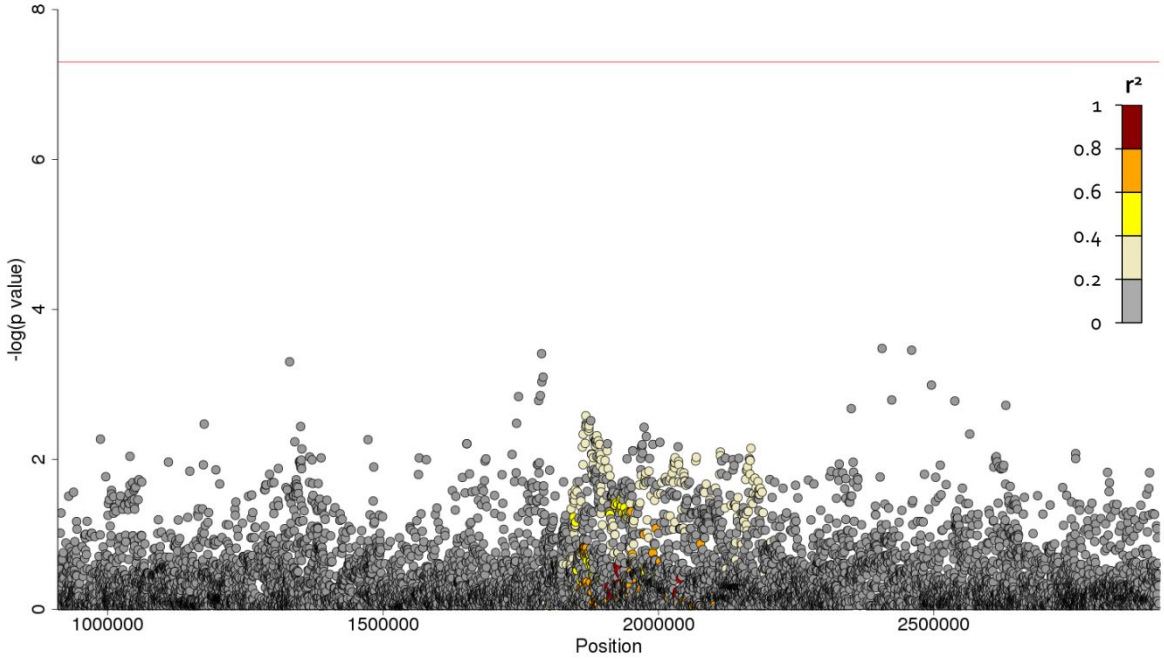
AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

ii) Chromosome 7

Region plot

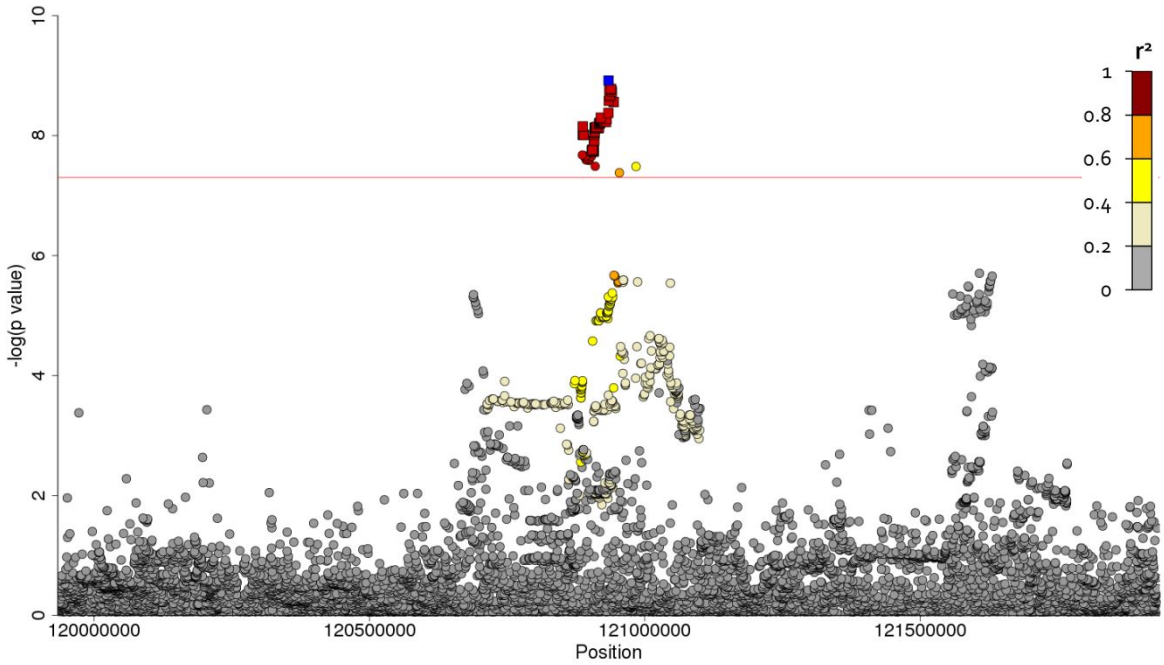


After conditioning on sentinel variant

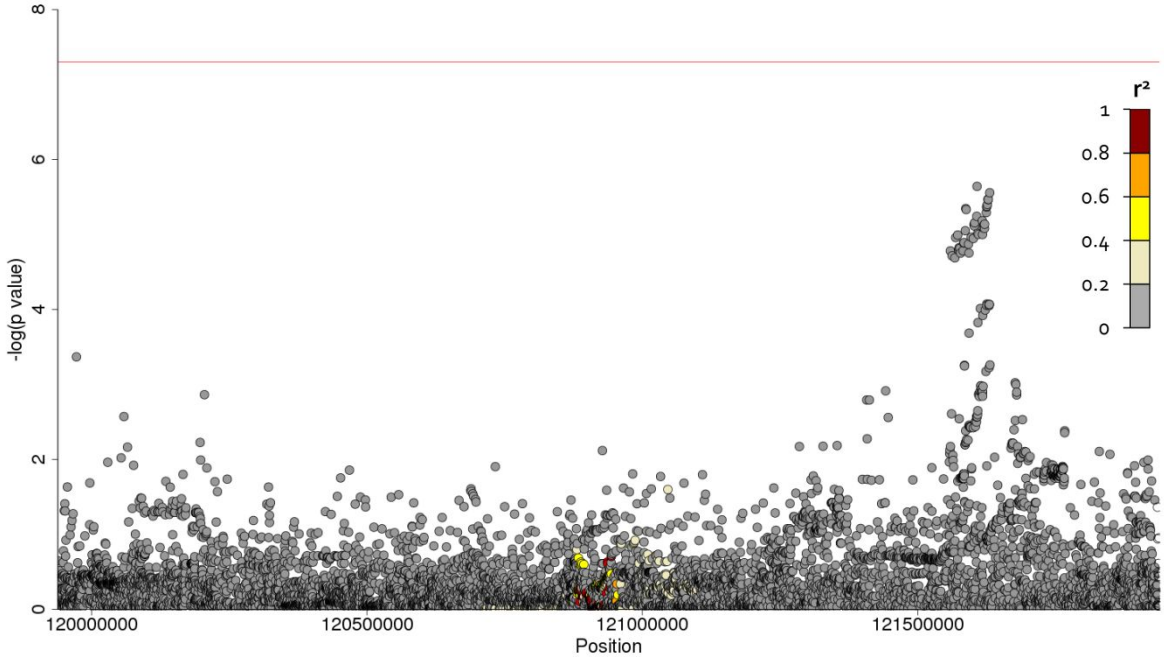


iii) Chromosome 8

Region plot



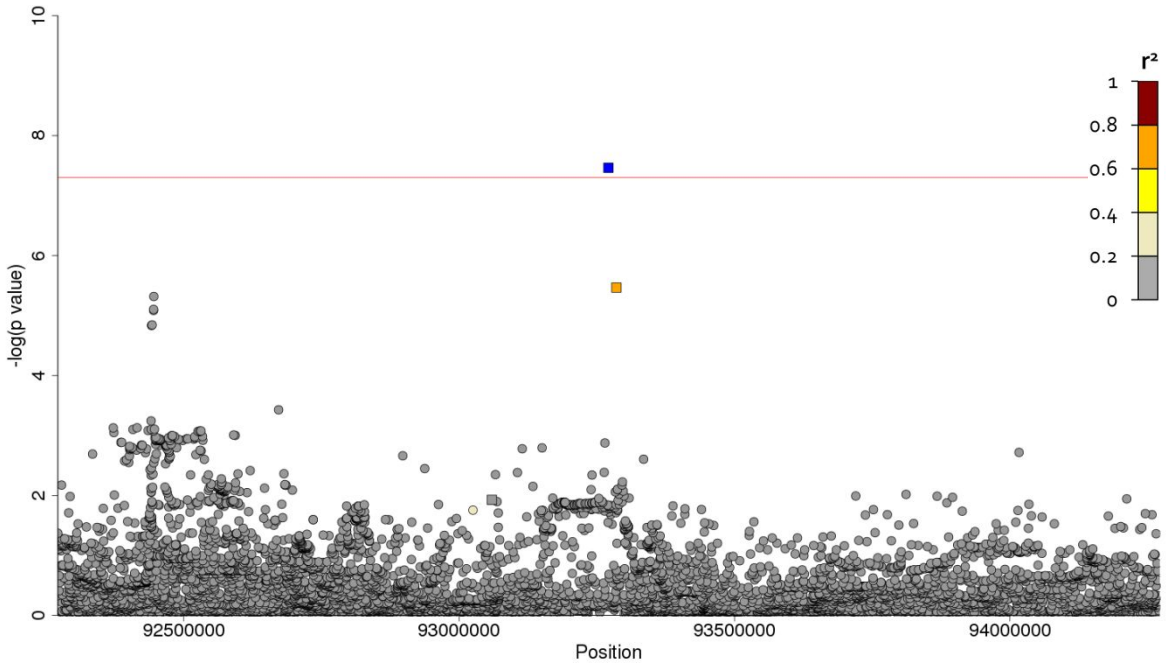
After conditioning on sentinel variant



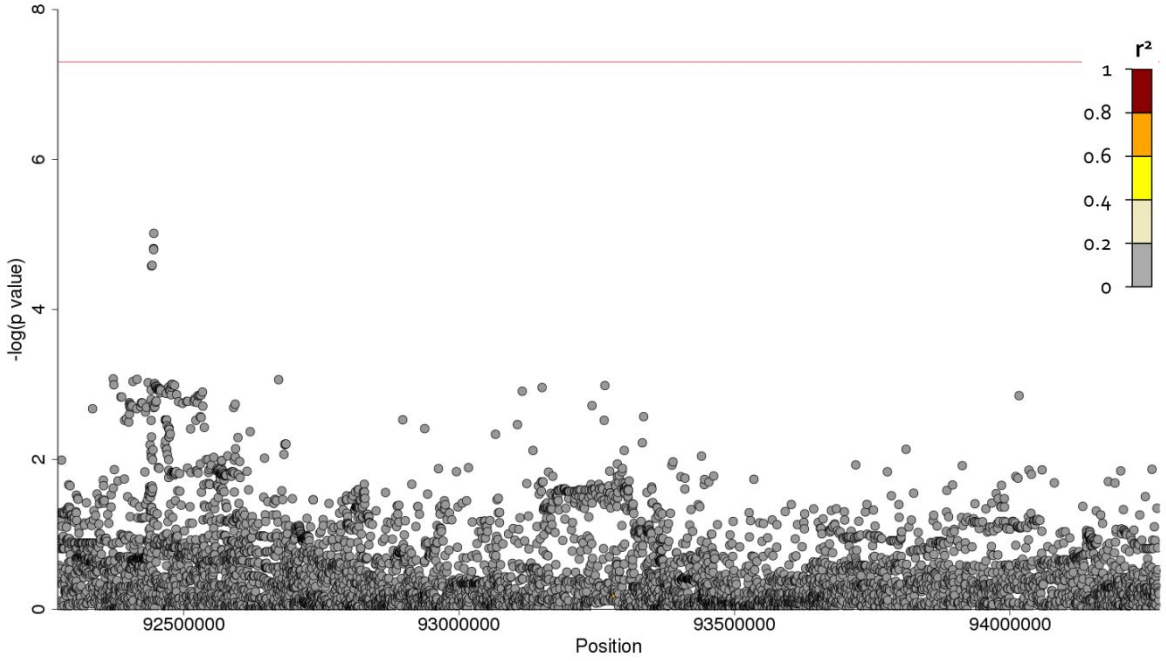
AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

iv) Chromosome 10

Region plot

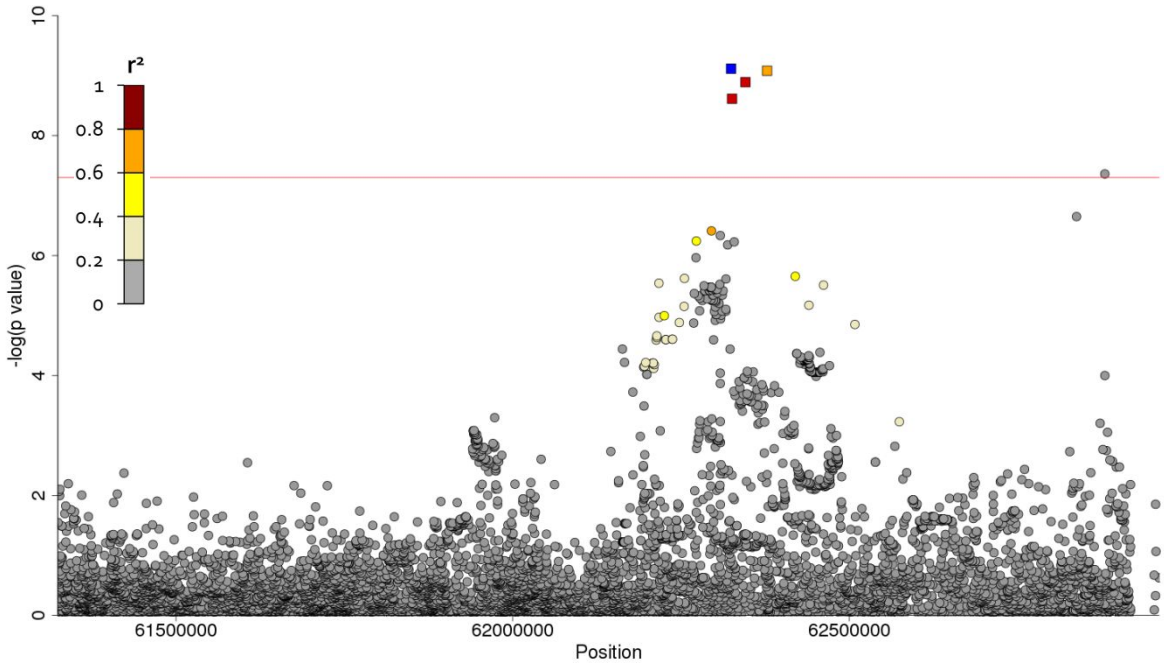


After conditioning on sentinel variant

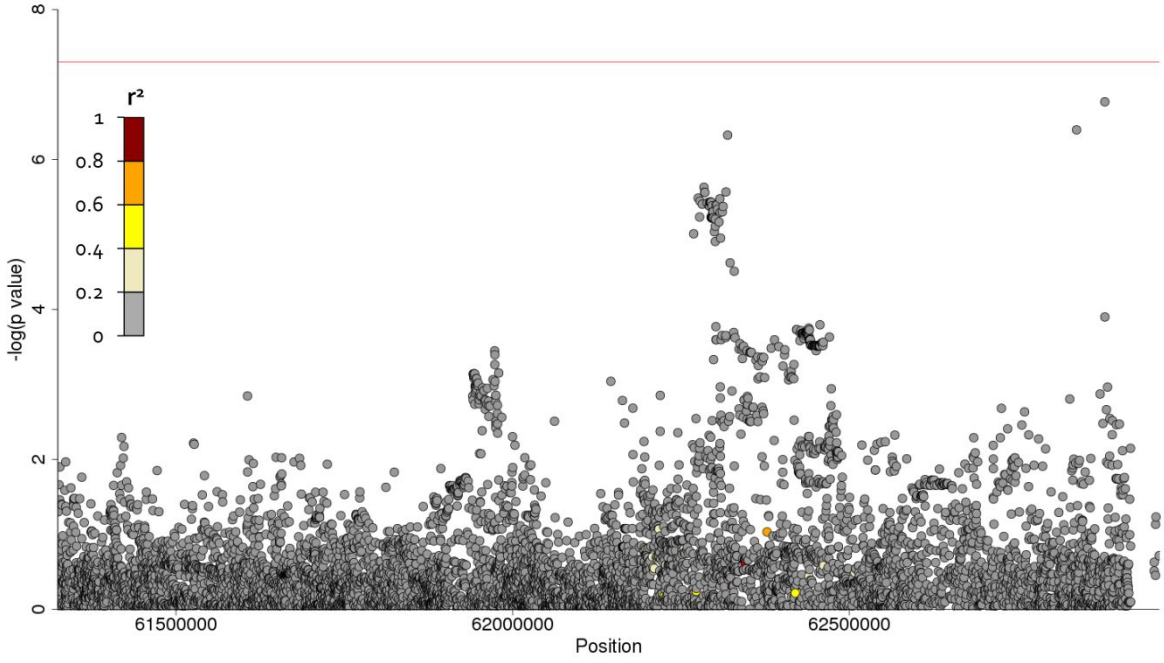


v) Chromosome 20

Region plot



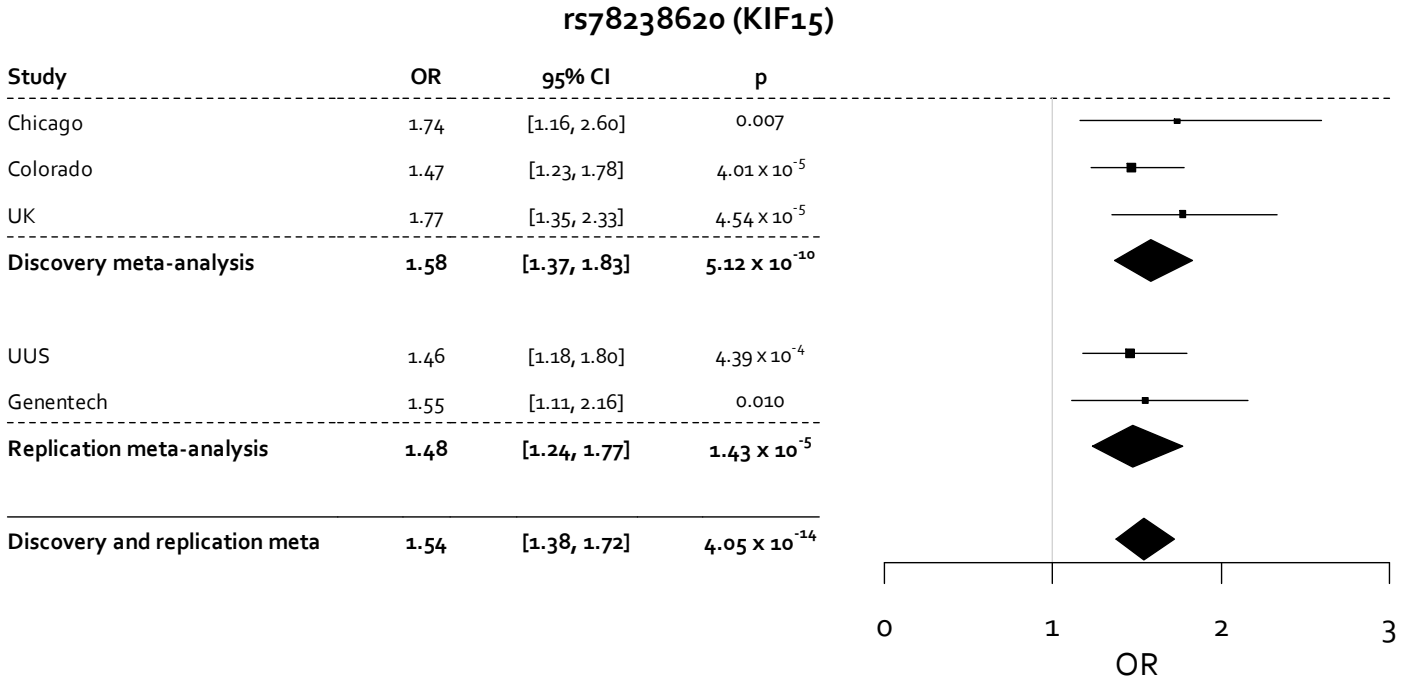
After conditioning on sentinel variant



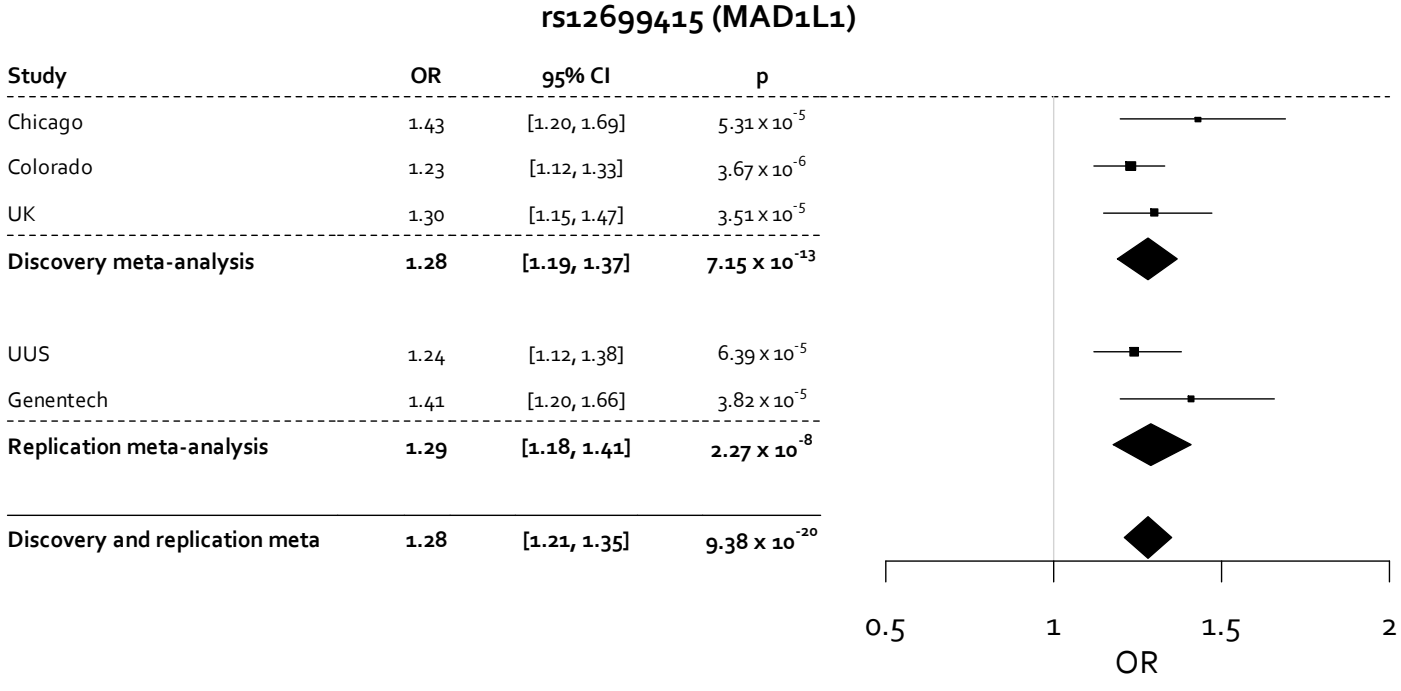
AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Figure E6 - Forest plot of discovery and replication study level results for the five not previously reported variants signals reaching genome-wide significance in the discovery meta-analysis

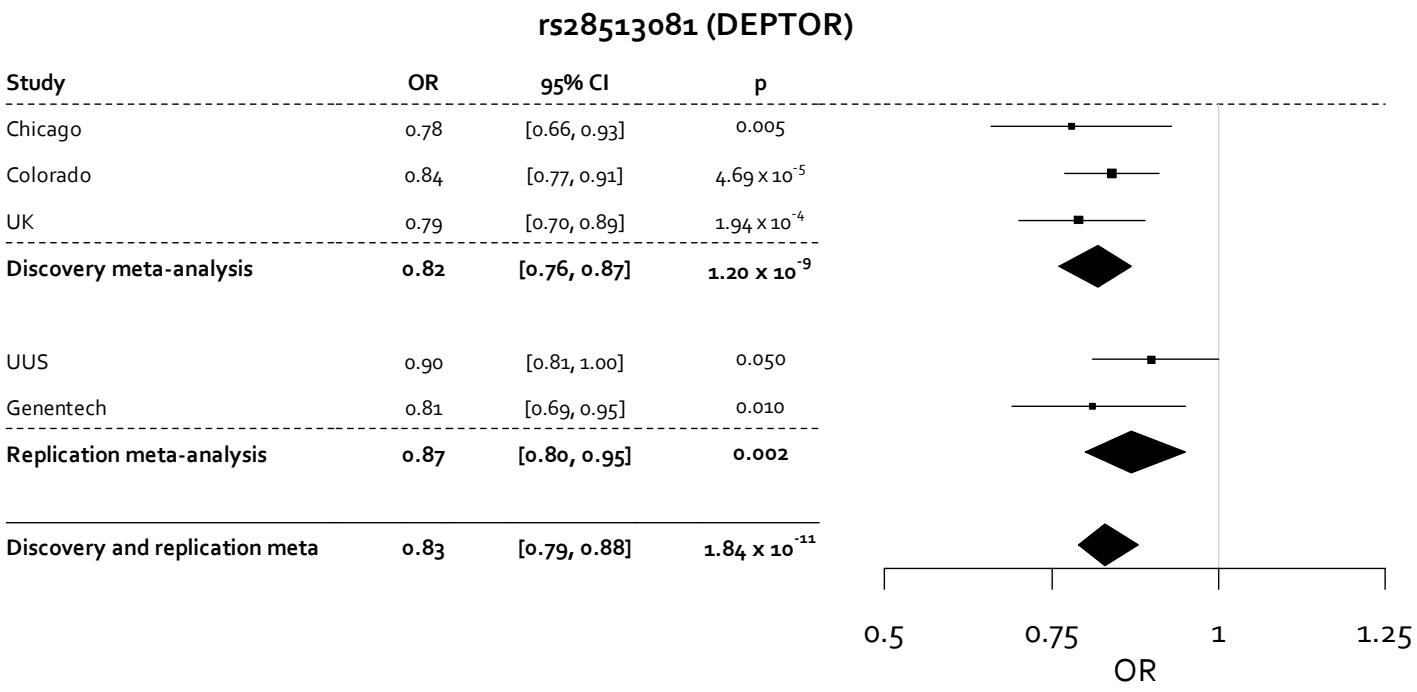
i) Forest plot for rs78238620



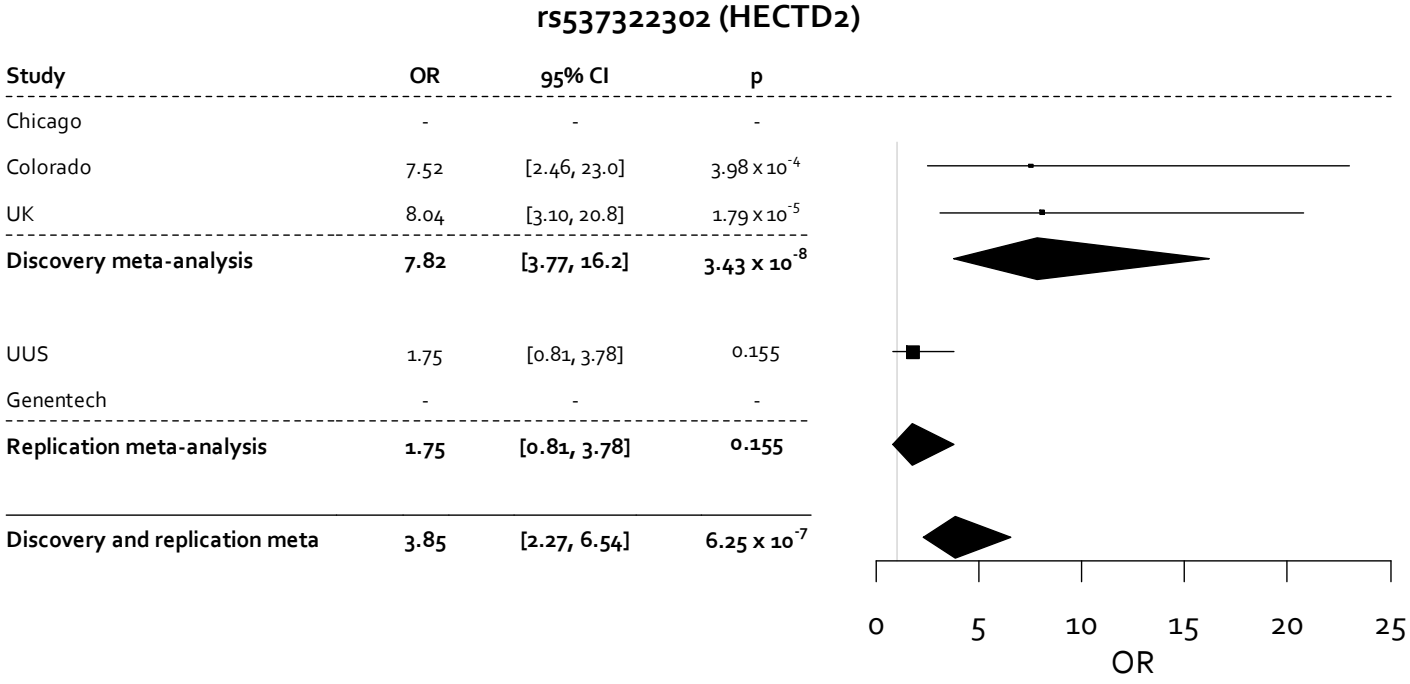
ii) Forest plot for rs12699415



iii) Forest plot for rs28513081



iv) Forest plot for rs537322302



AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

v) Forest plot for rs41308092

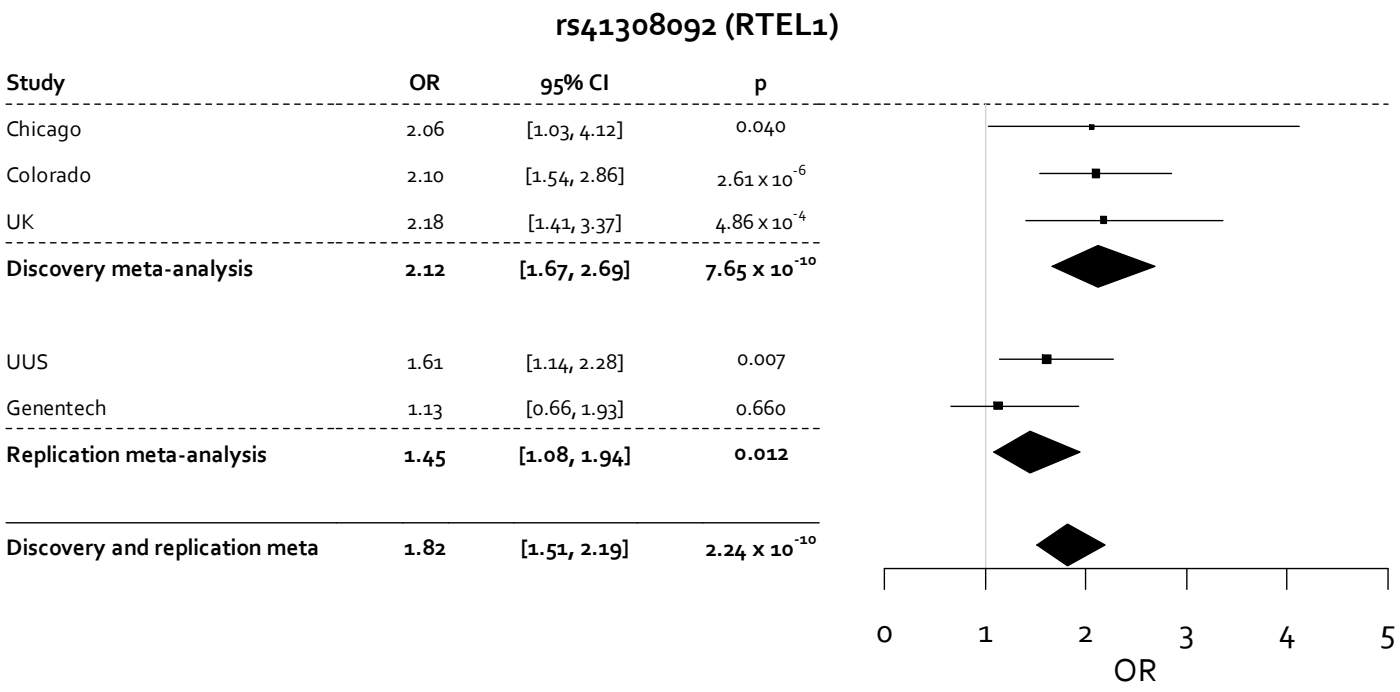
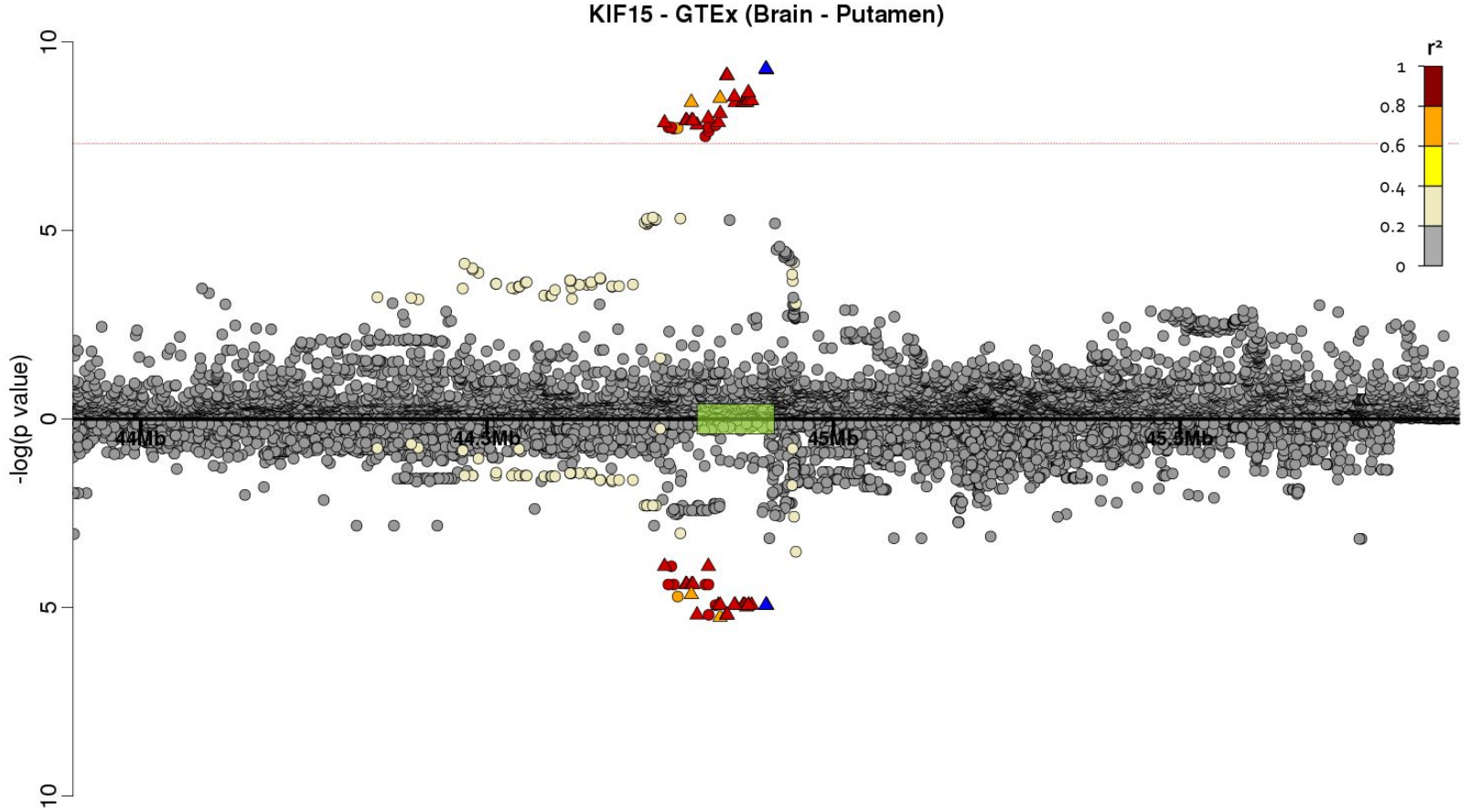


Figure E7 - GWAS vs eQTL results for novel IPF susceptibility signal on chromosome 3

Each point represents a variant with chromosomal position on the x axis and −log(*P* value) on the y axis. Above the x axis is the −log(*P* value) from the IPF susceptibility discovery genome-wide meta-analysis and below the y axis is the −log(*P* value) from the eQTL database. The sentinel variant from the IPF susceptibility analysis is coloured in blue with all other variants coloured by LD with the sentinel variant (variants in red have $r^2 \geq 0.8$ with the sentinel variant, variants in orange have $0.6 \leq r^2 < 0.8$, variants in yellow have $0.4 \leq r^2 < 0.6$, variants in light yellow have $0.2 \leq r^2 < 0.4$ and variants in grey have $r^2 < 0.2$ with the sentinel variant. The area in green on the x axis denotes the location of the gene implicated by the eQTL analysis.

i) *KIF15* - Brain (Putamen) - Colocalisation probability = 95.6%



AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

ii) *TMEM42* - Thyroid - Colocalisation probability = 93.1%

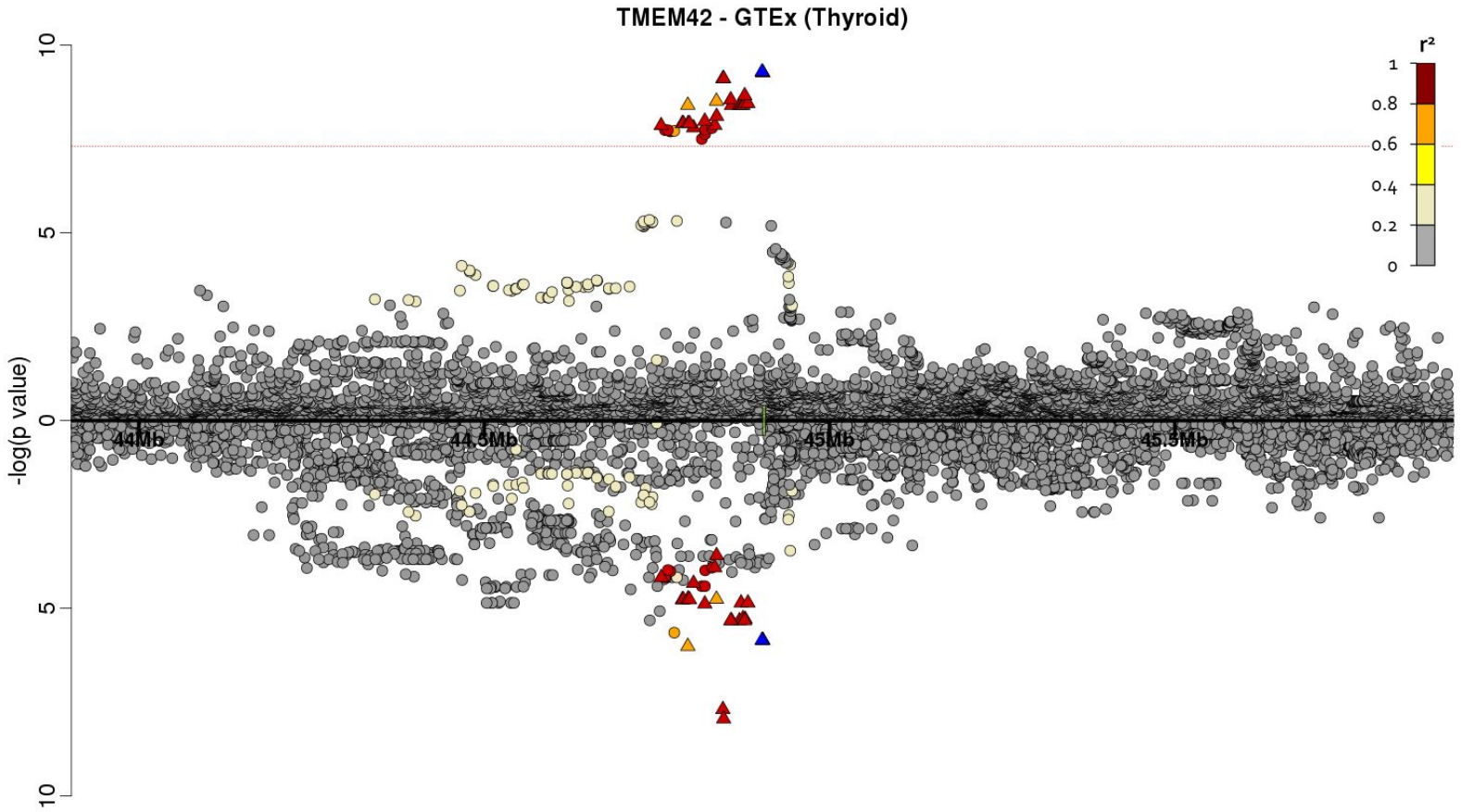
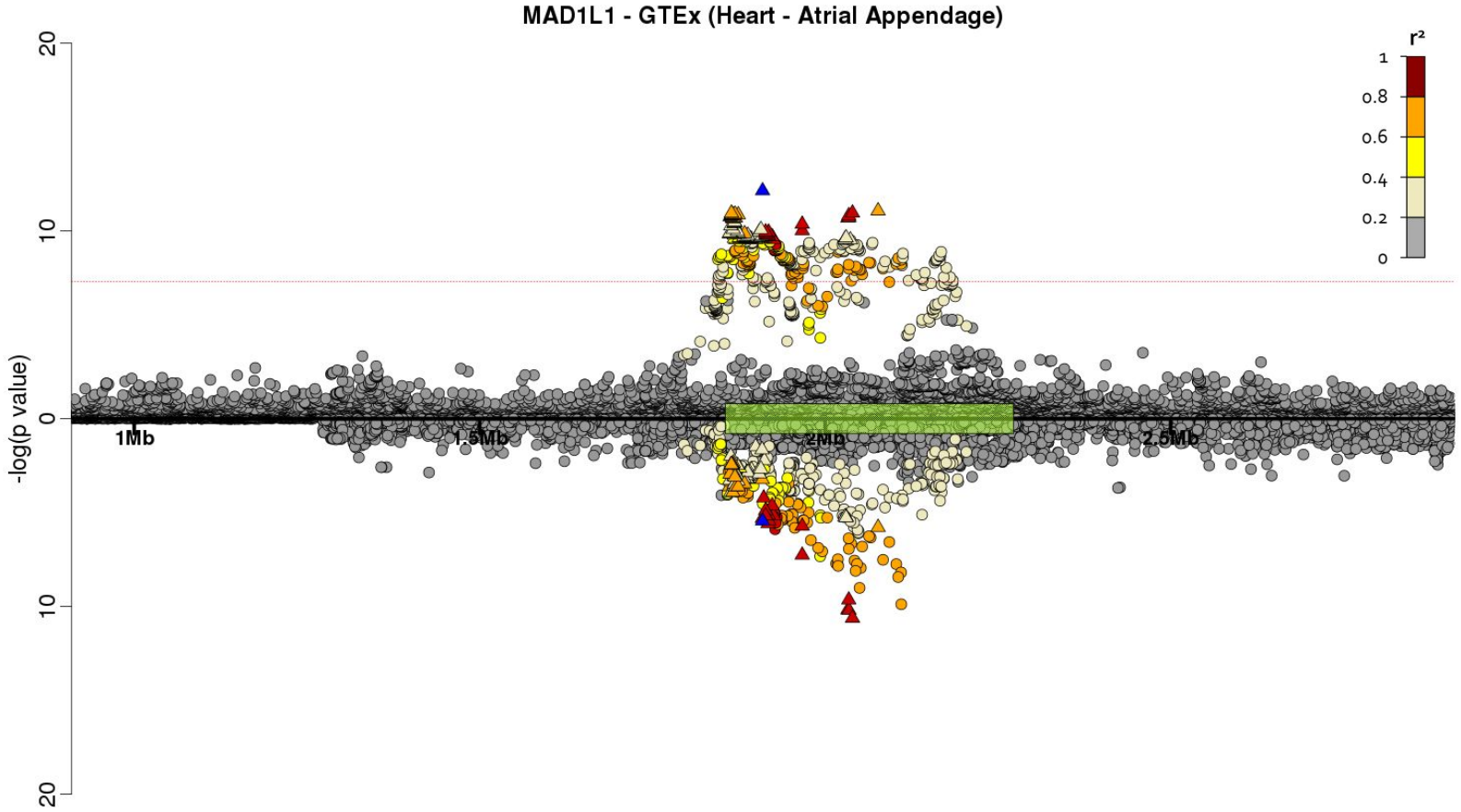


Figure E8 - GWAS vs eQTL for novel IPF susceptibility signal on chromosome 7

Each point represents a variant with chromosomal position on the x axis and −log(*P* value) on the y axis. Above the x axis is the −log(*P* value) from the IPF susceptibility discovery genome-wide meta-analysis and below the y axis is the −log(*P* value) from the eQTL database. The sentinel variant from the IPF susceptibility analysis is coloured in blue with all other variants coloured by LD with the sentinel variant (variants in red have $r^2 \geq 0.8$ with the sentinel variant, variants in orange have $0.6 \leq r^2 < 0.8$, variants in yellow have $0.4 \leq r^2 < 0.6$, variants in light yellow have $0.2 \leq r^2 < 0.4$ and variants in grey have $r^2 < 0.2$ with the sentinel variant. The area in green on the x axis denotes the location of the gene implicated by the eQTL analysis.

i) *MAD1L1* - Heart (Atrial Appendage) - Colocalisation probability = 95.3%

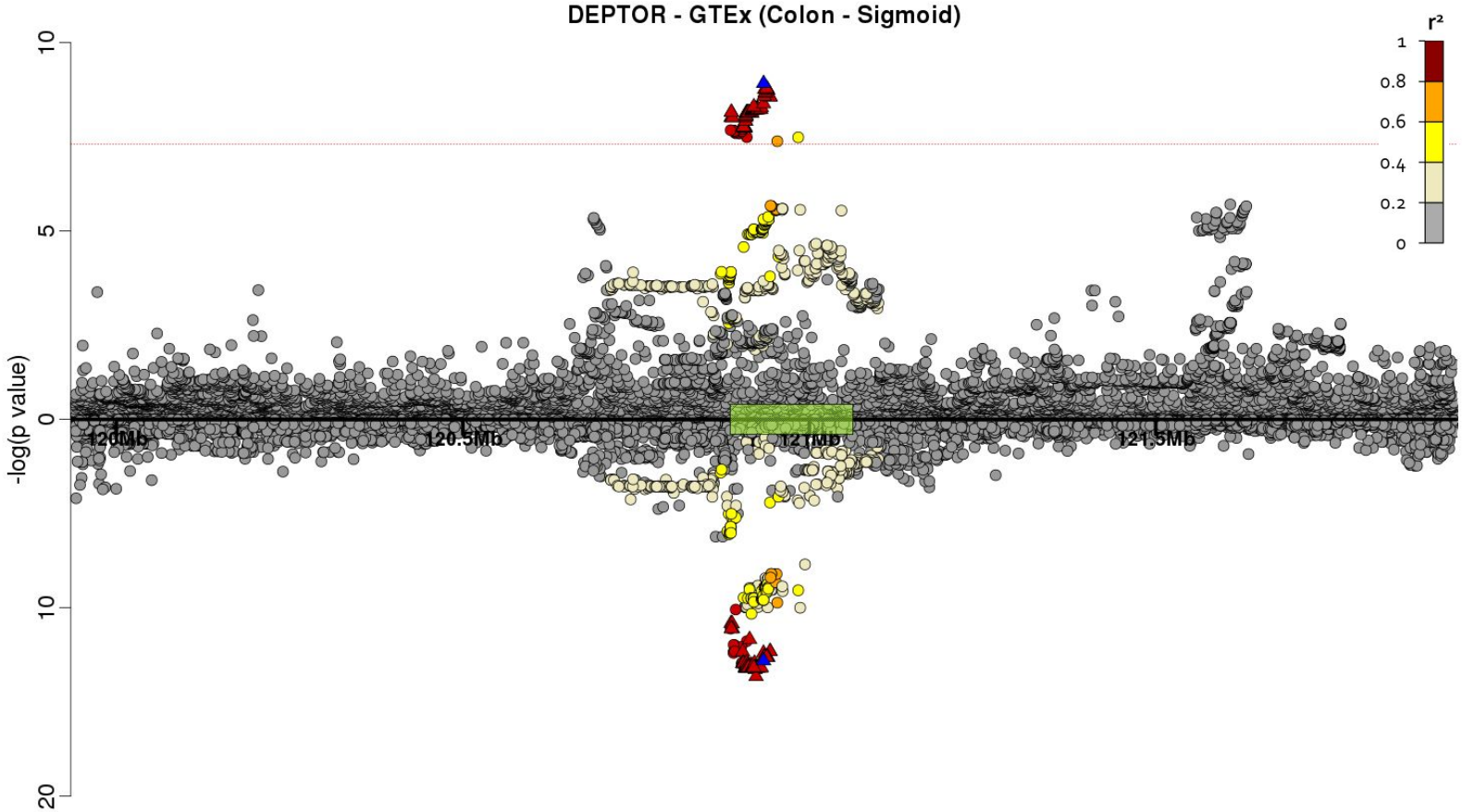


AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

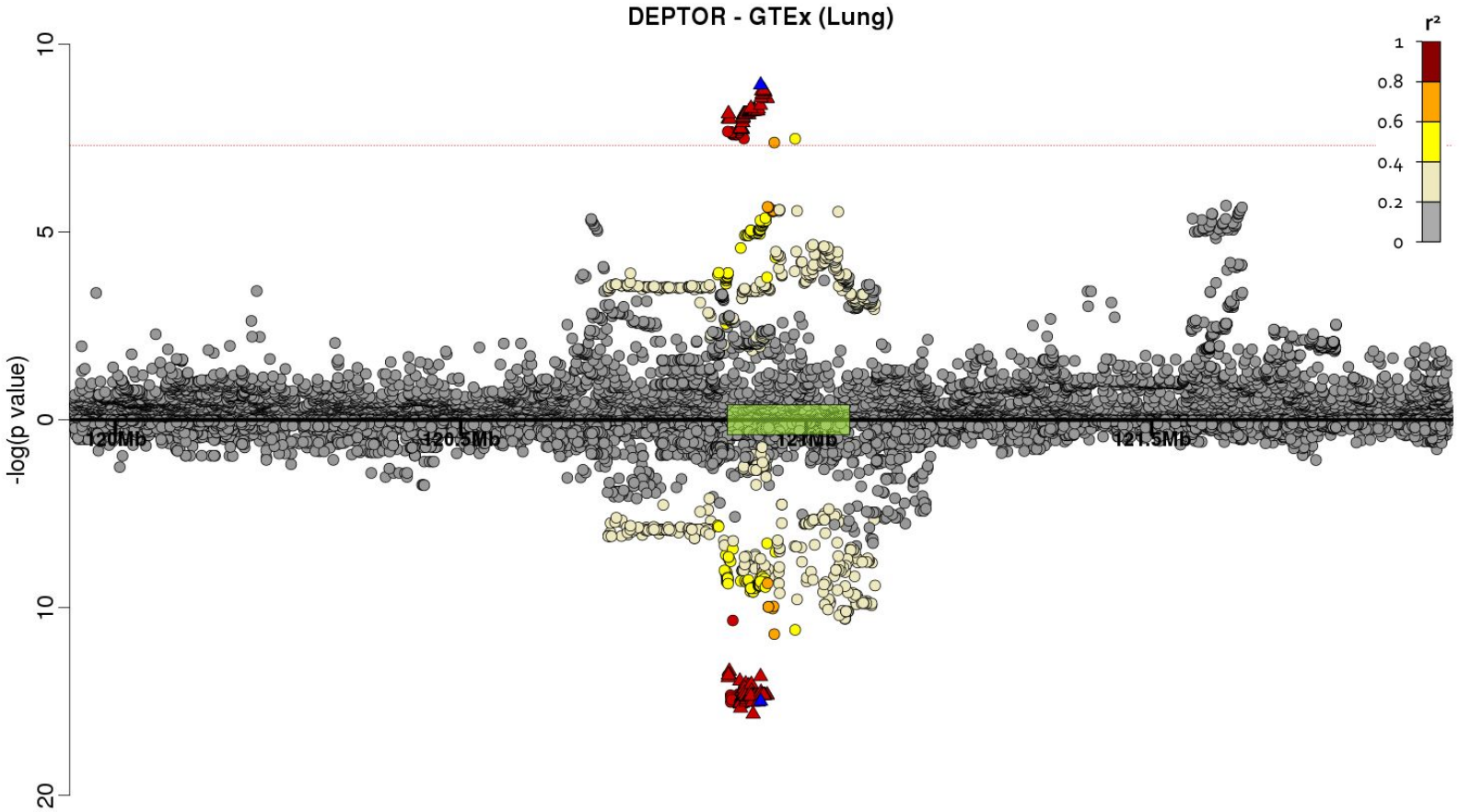
Figure E9 - GWAS vs eQTL for novel IPF susceptibility signal on chromosome 8

Each point represents a variant with chromosomal position on the x axis and −log(*P* value) on the y axis. Above the x axis is the −log(*P* value) from the IPF susceptibility discovery genome-wide meta-analysis and below the y axis is the −log(*P* value) from the eQTL database. The sentinel variant from the IPF susceptibility analysis is coloured in blue with all other variants coloured by LD with the sentinel variant (variants in red have $r^2 \geq 0.8$ with the sentinel variant, variants in orange have $0.6 \leq r^2 < 0.8$, variants in yellow have $0.4 \leq r^2 < 0.6$, variants in light yellow have $0.2 \leq r^2 < 0.4$ and variants in grey have $r^2 < 0.2$ with the sentinel variant. The area in green on the x axis denotes the location of the gene implicated by the eQTL analysis.

i) *DEPTOR* - Colon (Sigmoid) - Colocalisation probability = 89.6%

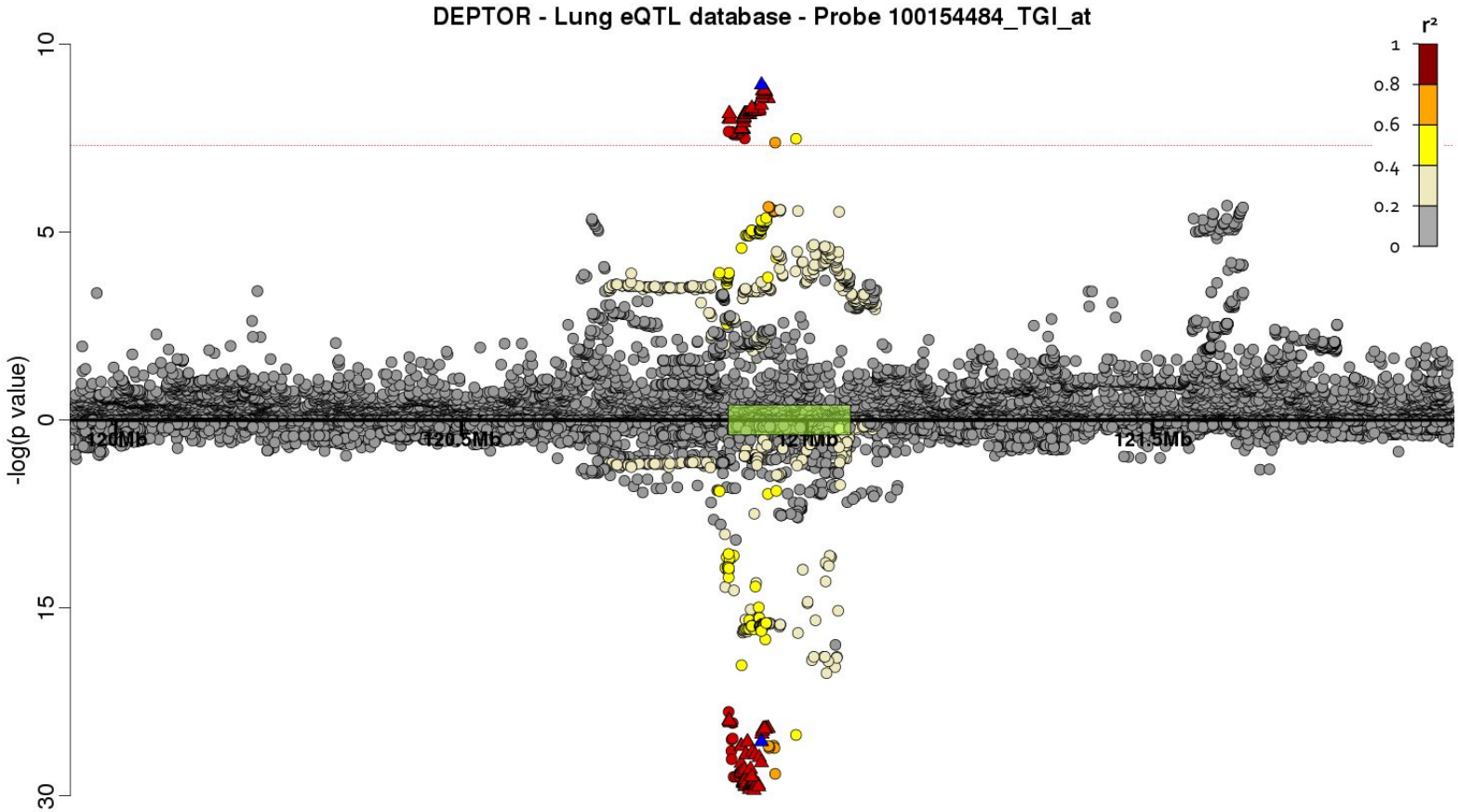


ii) *DEPTOR* - Lung - Colocalisation probability = 89.2%



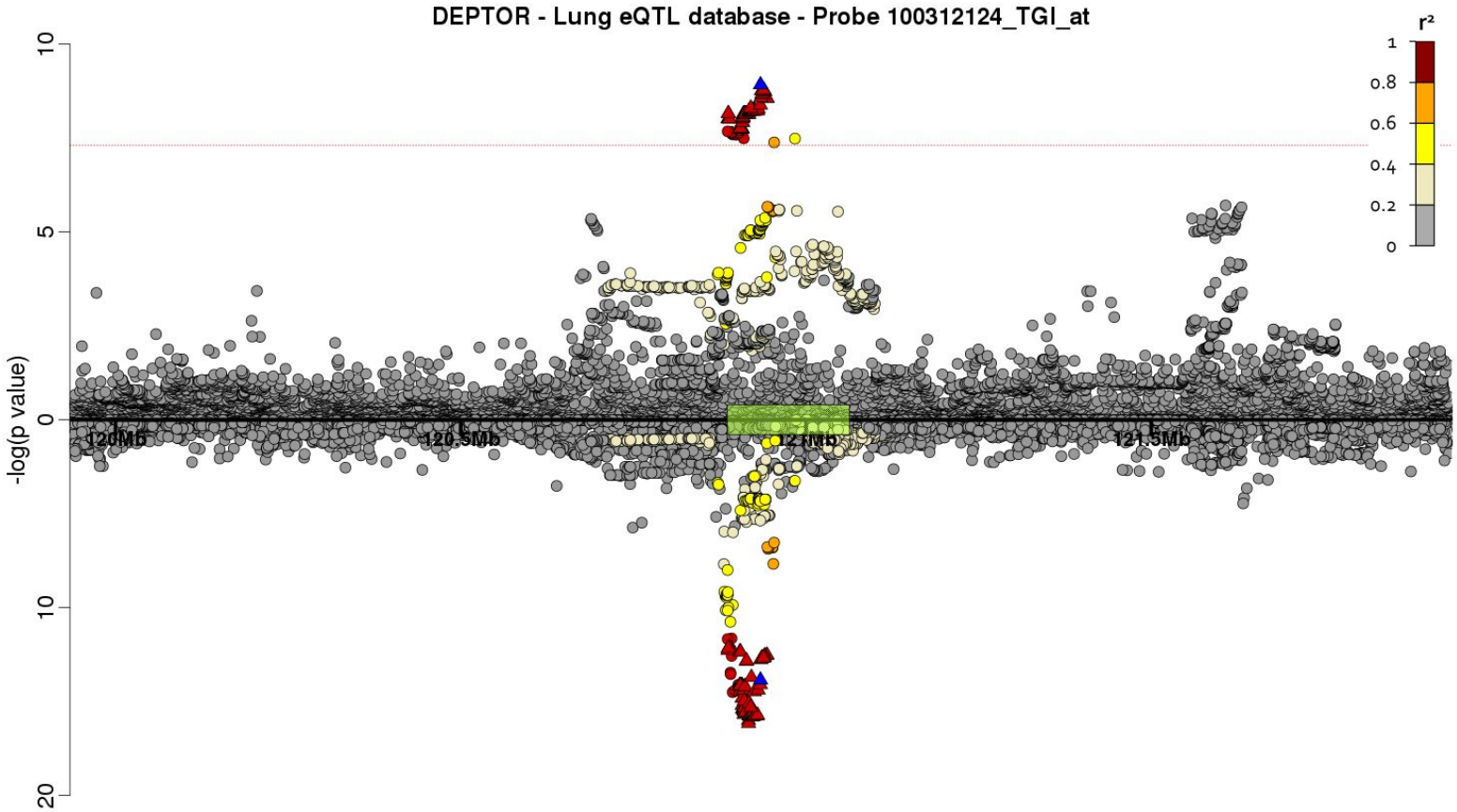
AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

iii) *DEPTOR* - Lung - Colocalisation probability = 89.5%



Note: The lung eQTL dataset showed two independent signals of association for *DEPTOR* expression. The eQTL results here are those obtained after conditioning on the top eQTL for *DEPTOR* to condition out the strongest signal which was driven by different variants to those driving the IPF risk association.

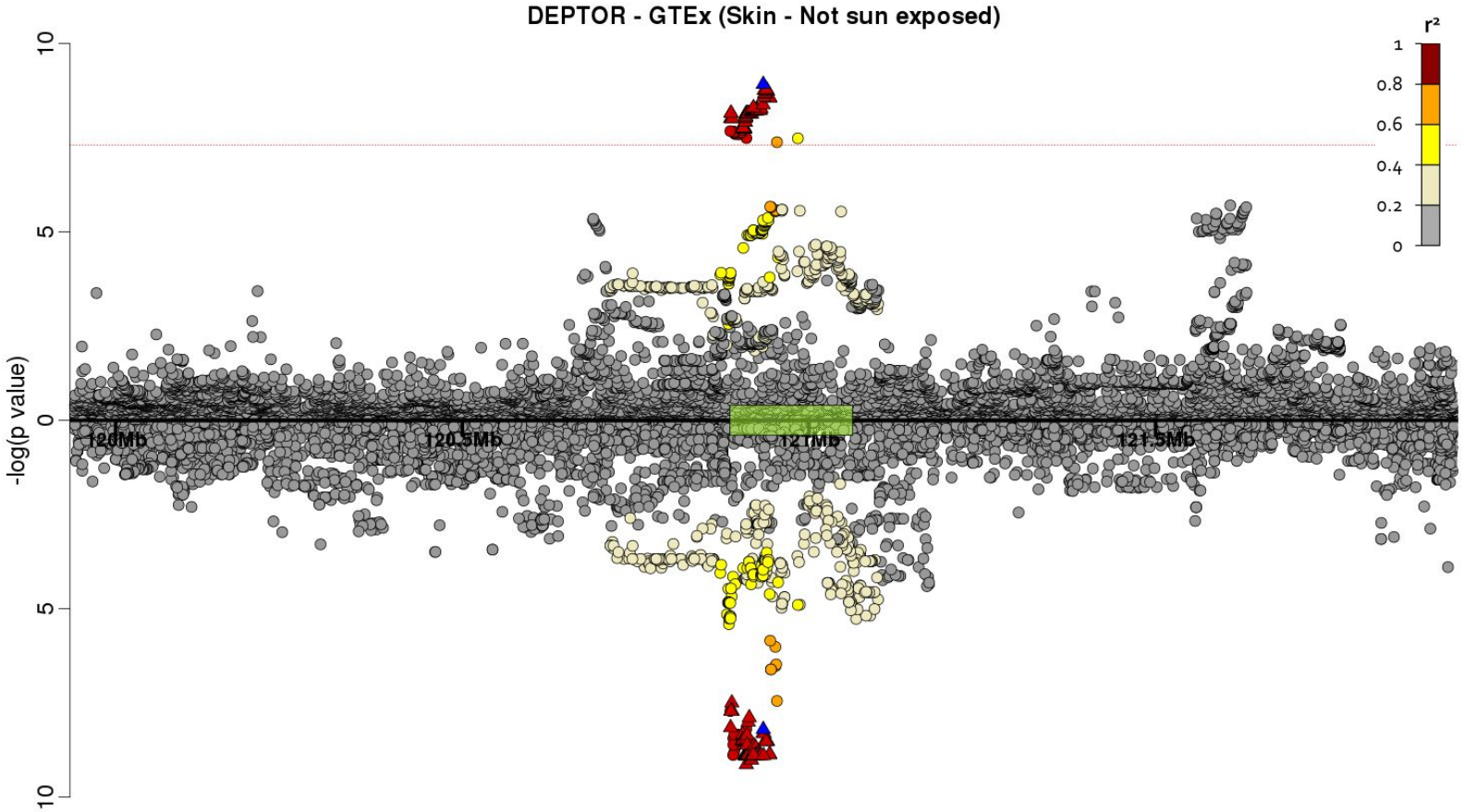
iv) *DEPTOR* - Lung - Colocalisation probability = 89.9%



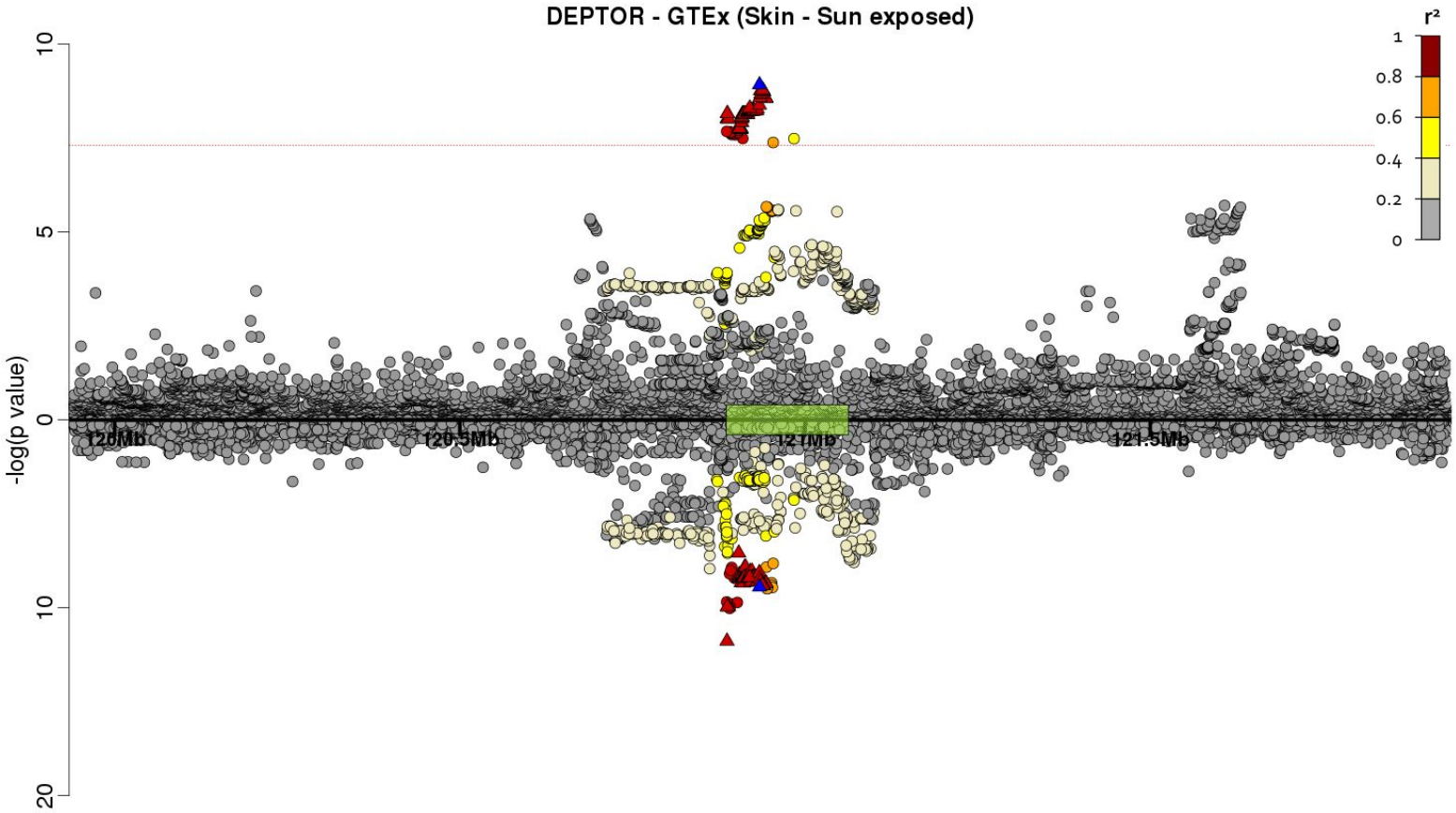
Note: The lung eQTL dataset showed two independent signals of association for *DEPTOR* expression. The eQTL results here are those obtained after conditioning on the top eQTL for *DEPTOR* to condition out the strongest signal which was driven by different variants to those driving the IPF risk association.

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

v) *DEPTOR* - Skin (Not sun exposed) - Colocalisation probability = 90.0%

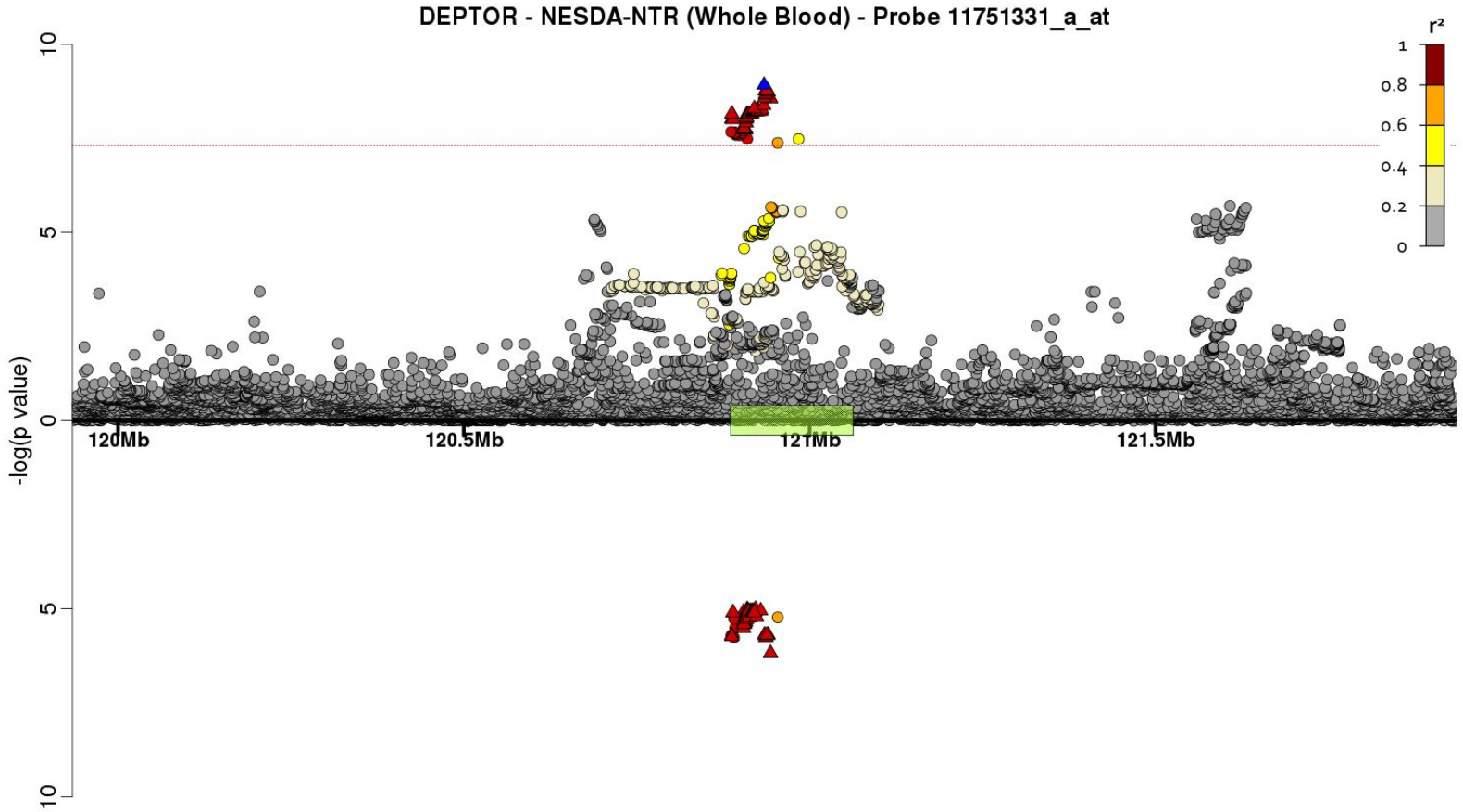


vi) *DEPTOR* - Skin (Sun exposed) - Colocalisation probability = 86.5%

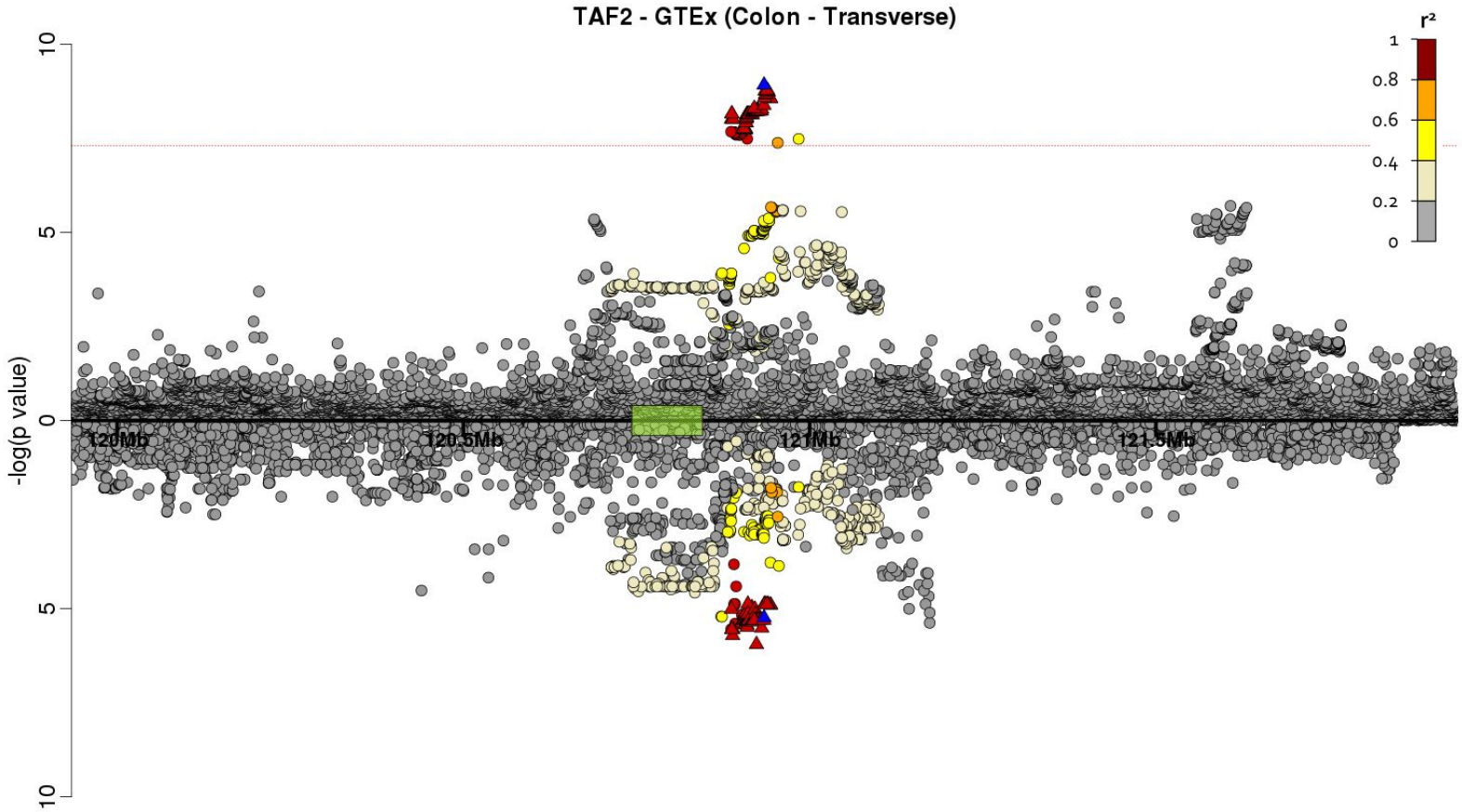


AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

vii) *DEPTOR* - Whole blood - Colocalisation probability = 93.7%

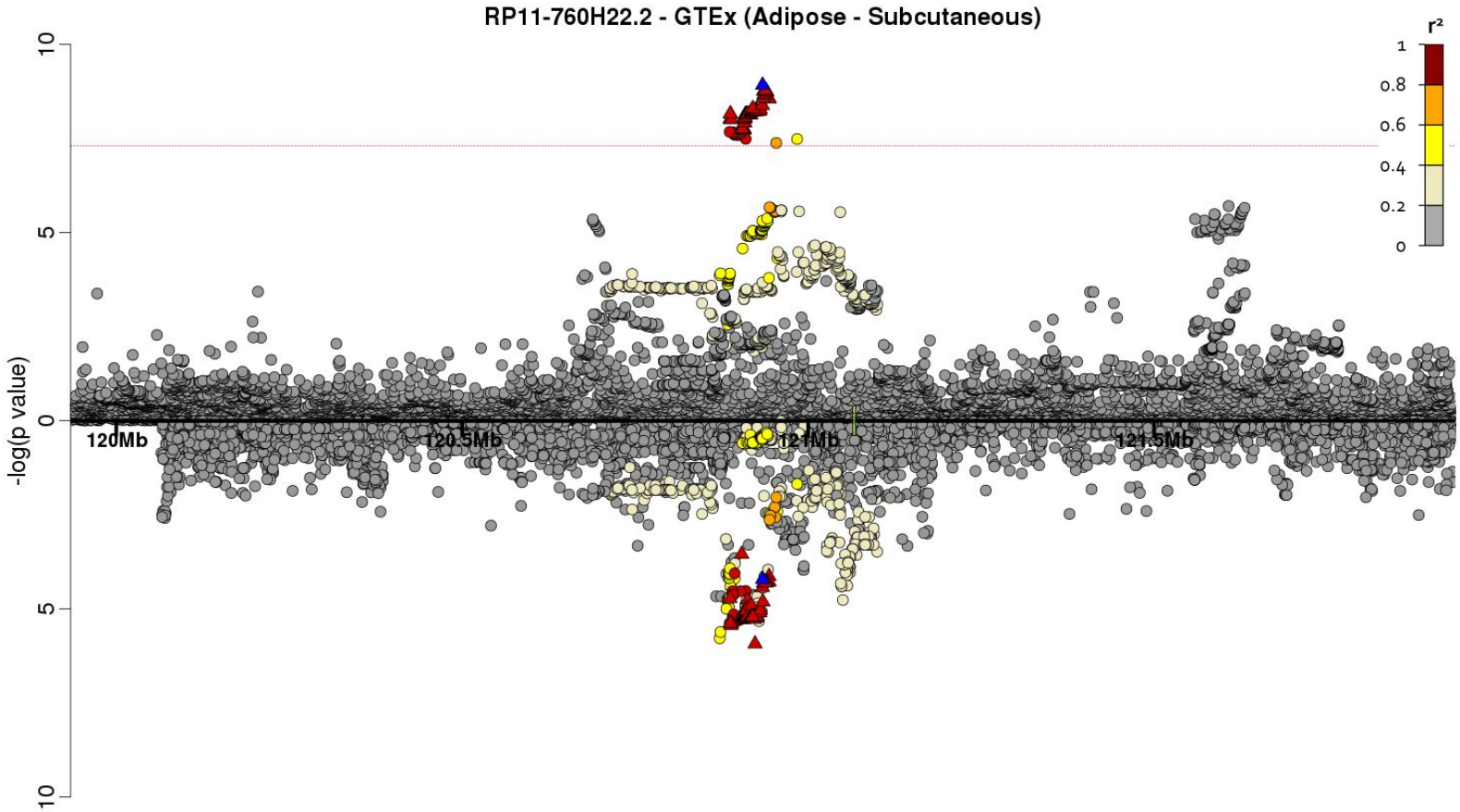


viii) *TAF2* - Colon (Transverse) - Colocalisation probability = 87.5%

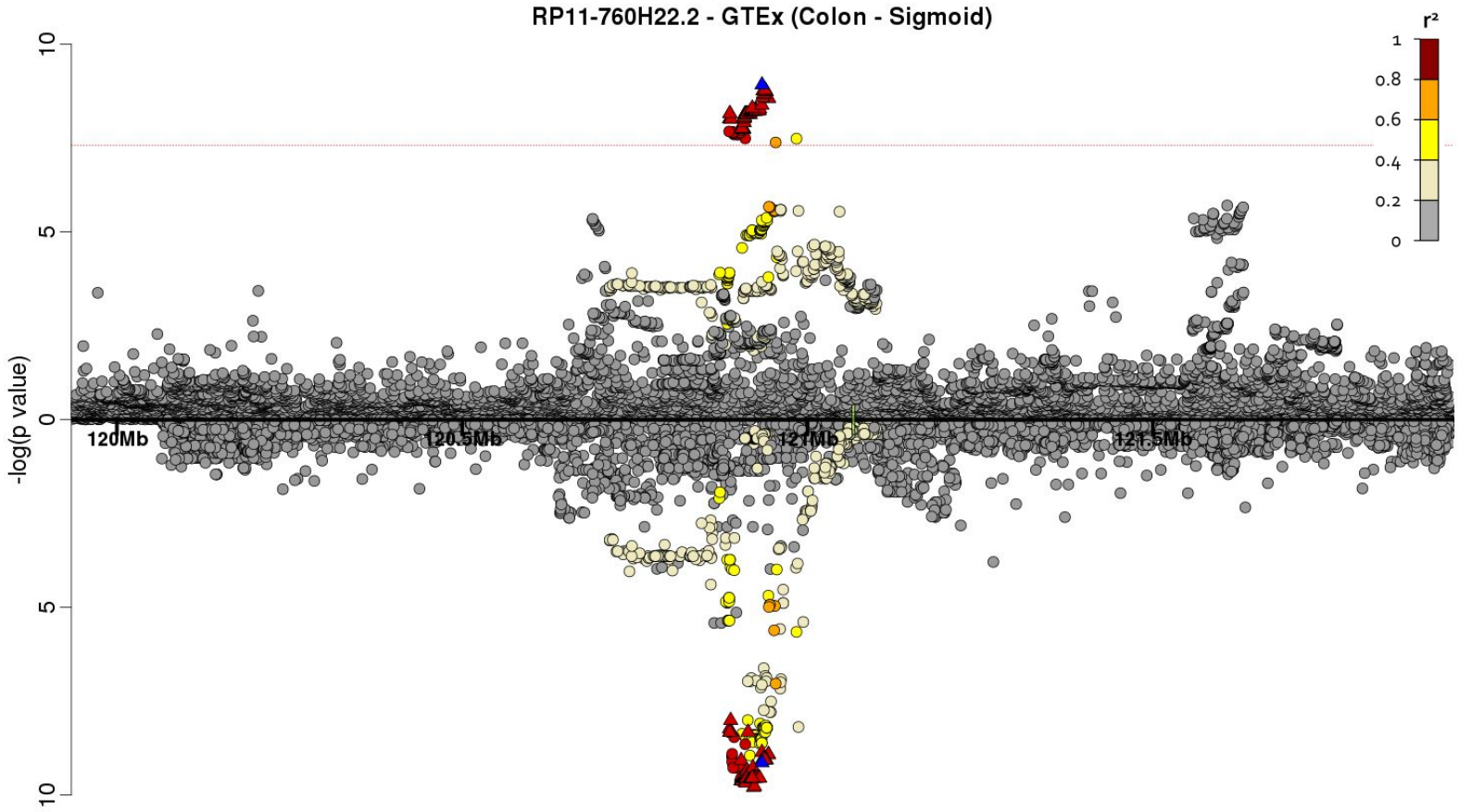


AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

ix) RP11-760H22.2 - Adipose (Subcutaneous) - Colocalisation probability = 84.9%

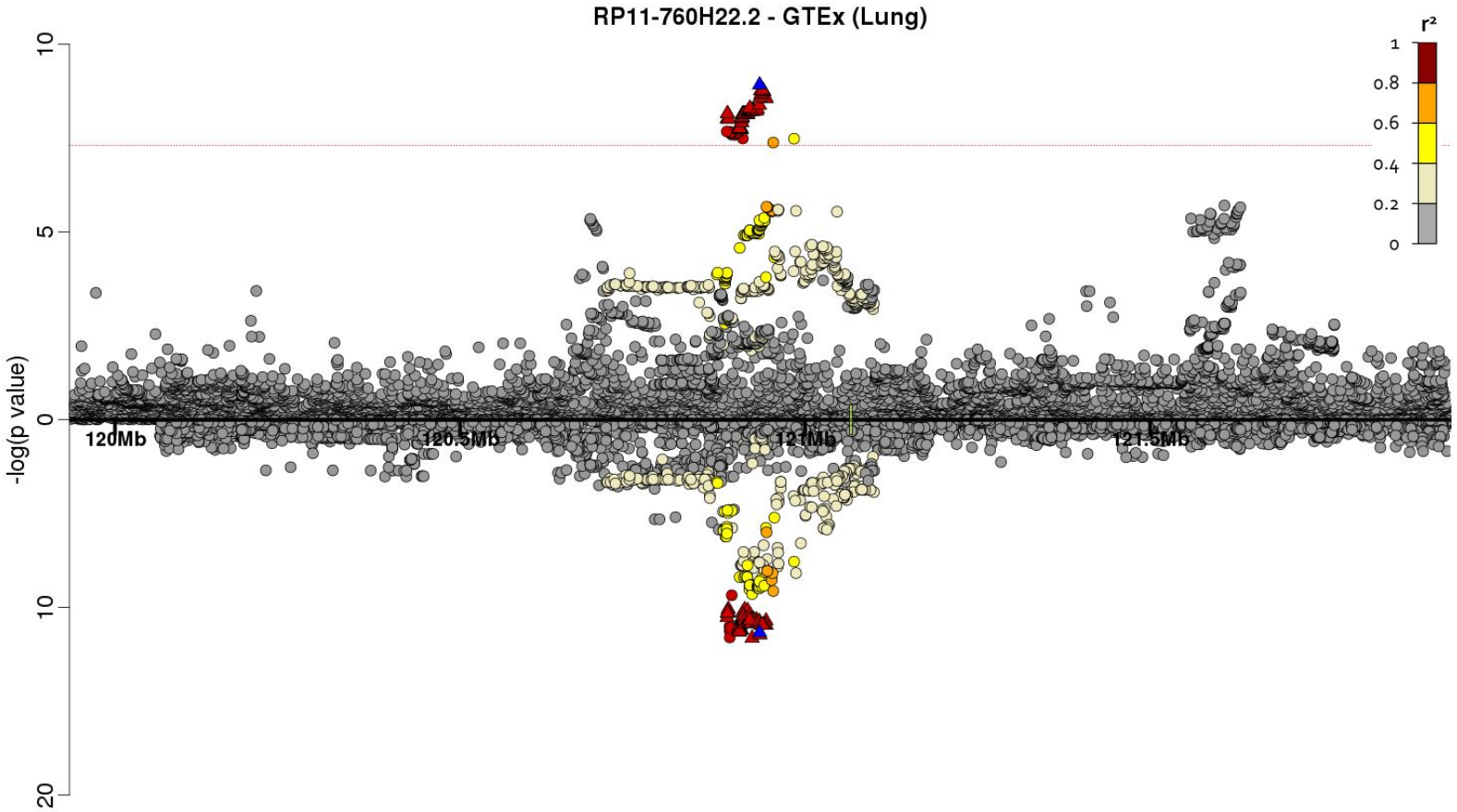


x) RP11-760H22.2 - Colon (Sigmoid) - Colocalisation probability = 88.6%

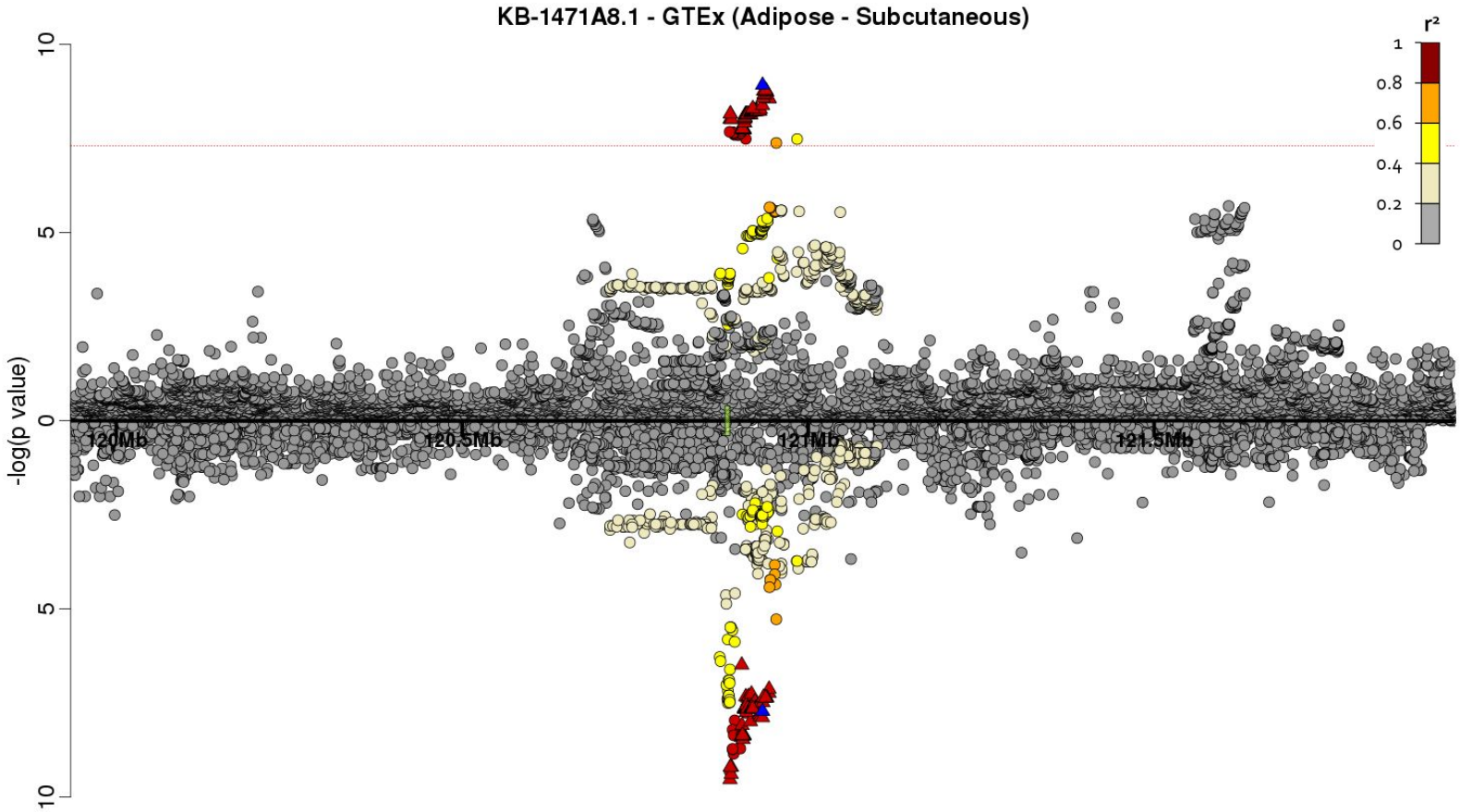


AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

xi) RP11-760H22.2 - Lung - Colocalisation probability = 90.0%

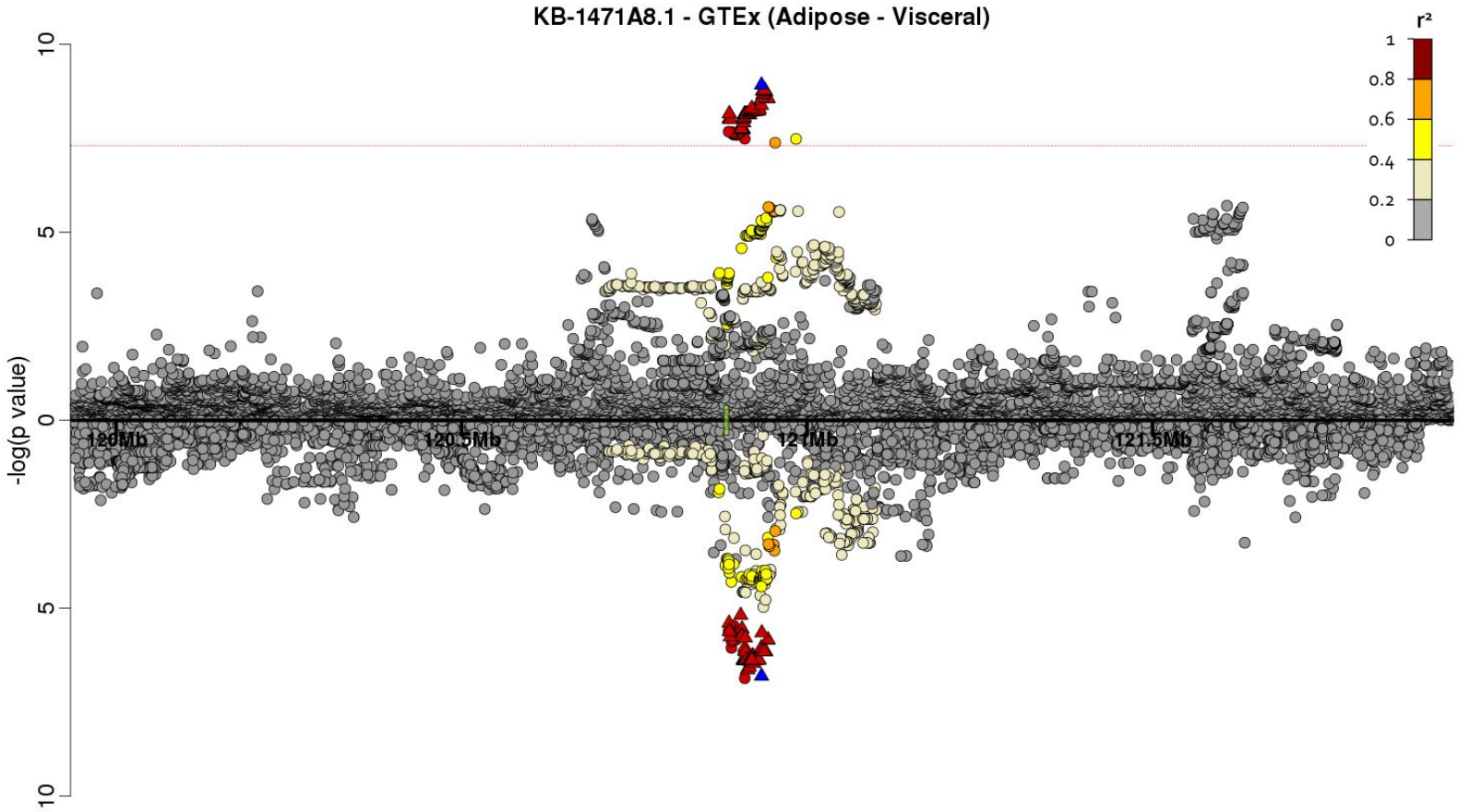


xii) KB-1471A8.1 - Adipose (Subcutaneous) - Colocalisation probability = 85.6%

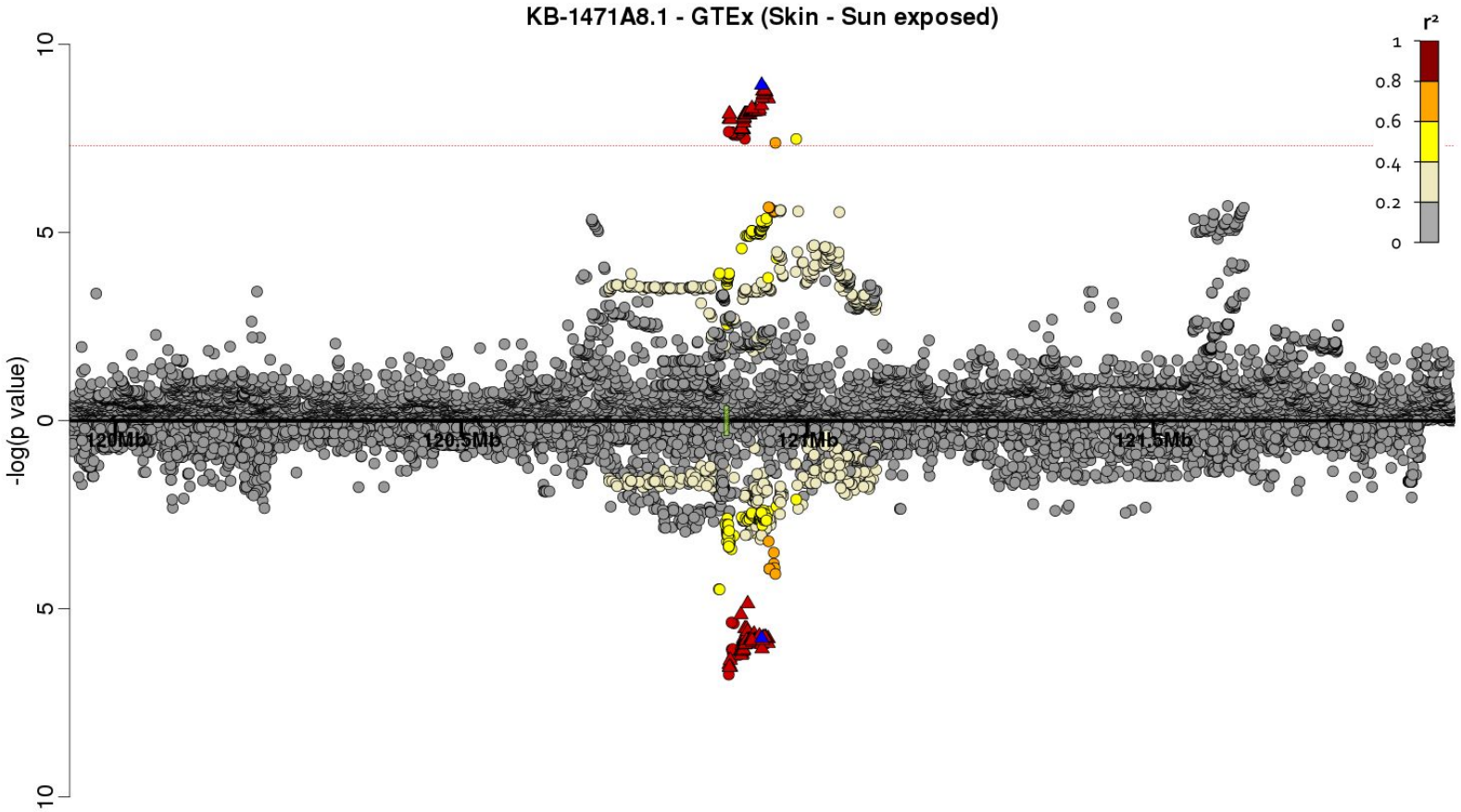


AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

xiii) KB-1471A8.1 - Adipose (Visceral) - Colocalisation probability = 90.9%



xiv) KB-1471A8.1 - Skin (Sun exposed) - Colocalisation probability = 88.7%



AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Figure E10 - FORGE analysis for enrichment of IPF susceptibility signals in regulatory regions

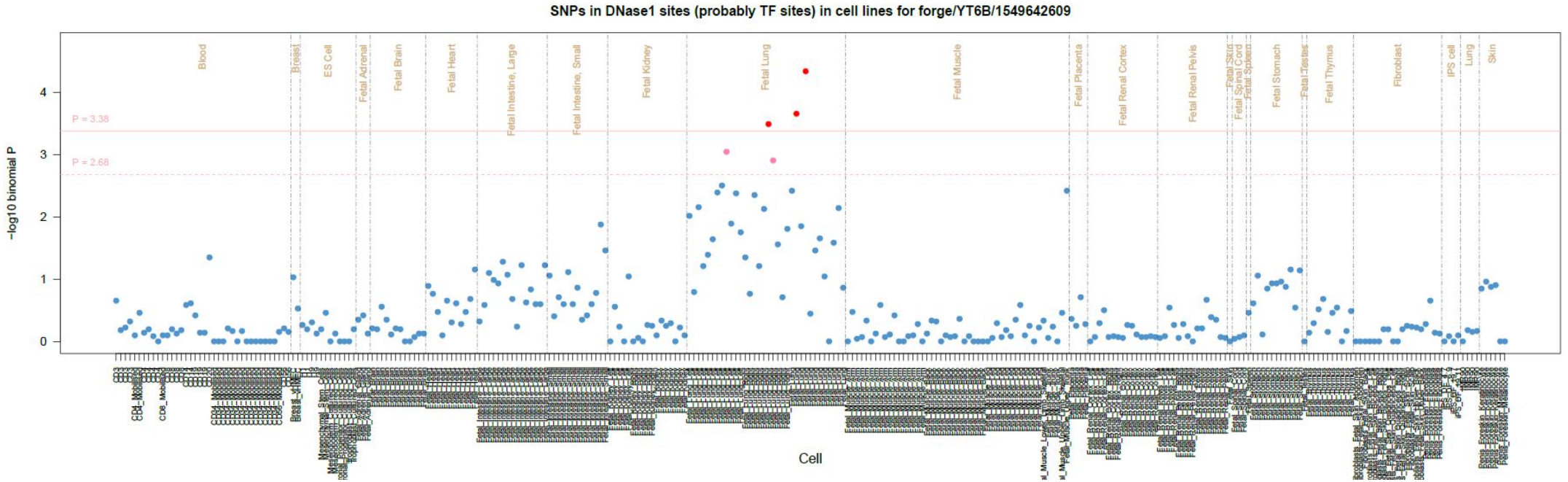
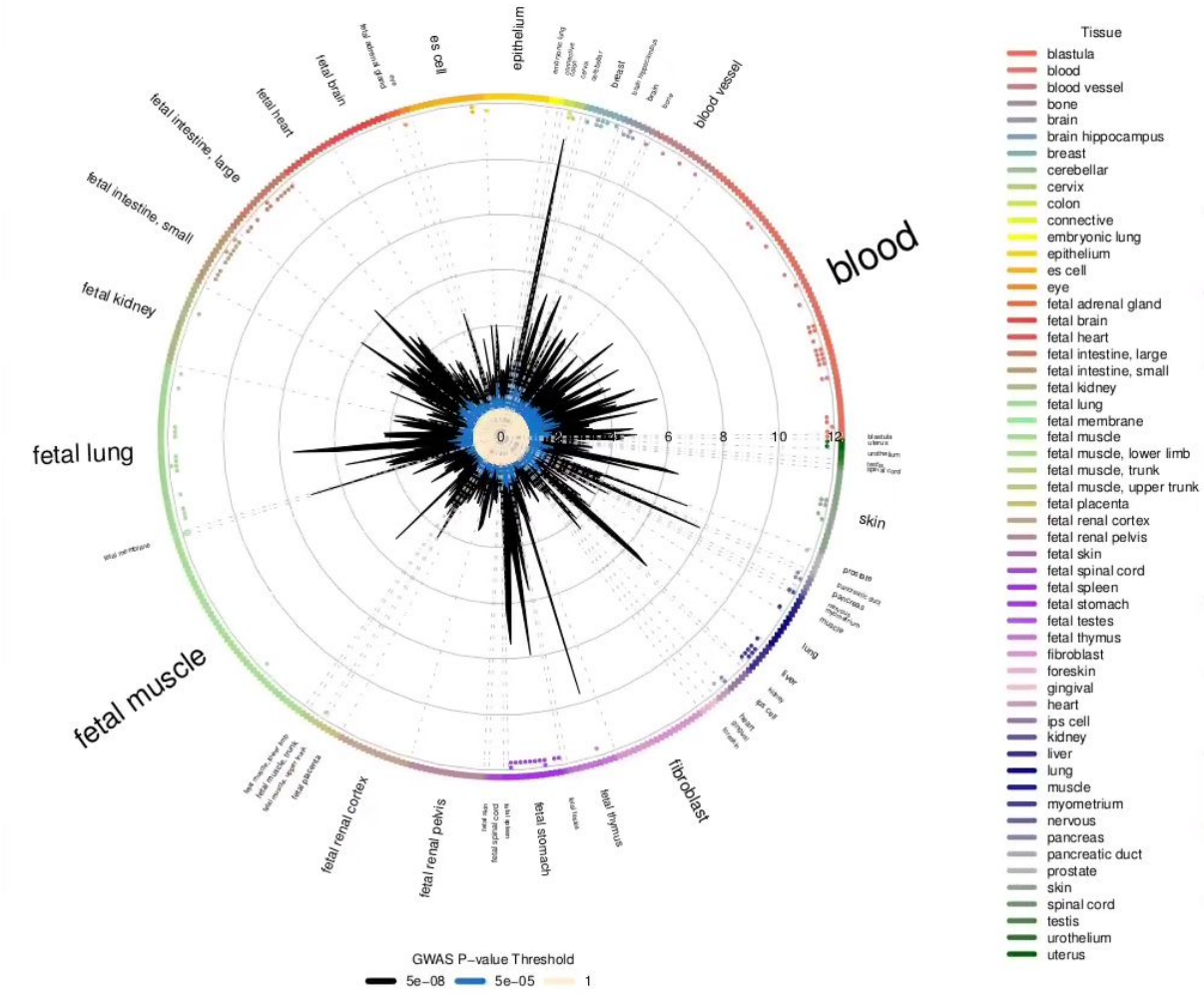


Figure E11 - GARFIELD analysis for enrichment of IPF susceptibility signals in DNase I hypersensitivity sites by tissue

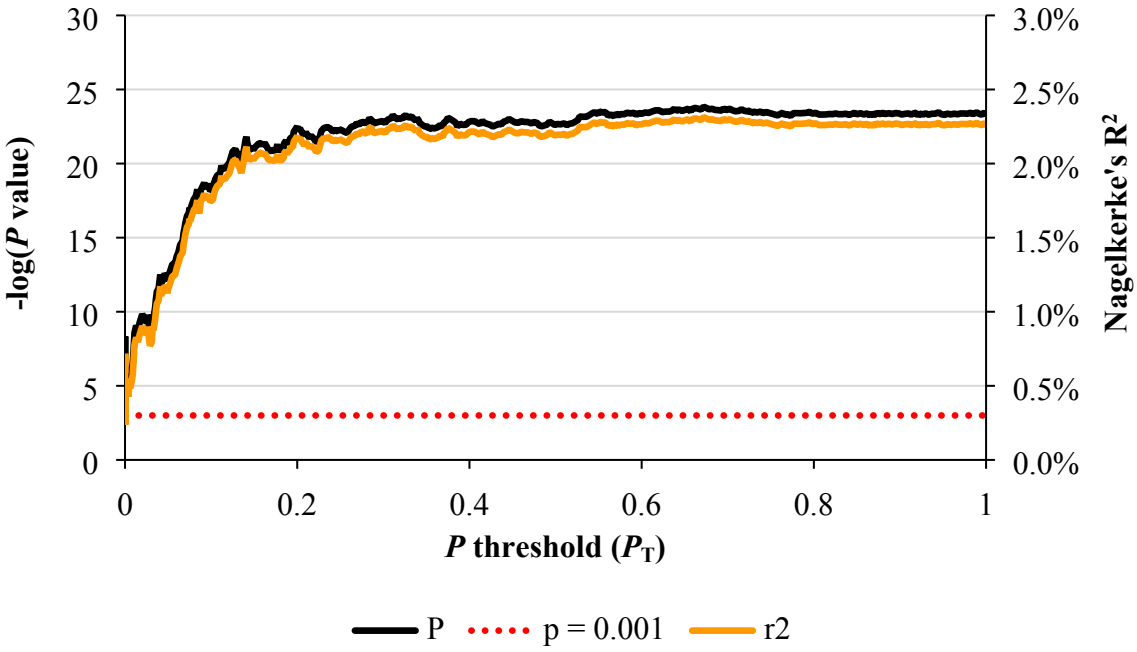
Radial plots for enrichment. The distance from the centre is equal to the odds ratio with the peaks shown in black to be when using a *P* threshold in the IPF genome-wide analysis of 5×10^{-8} and in blue for 5×10^{-5}. There are two rings of dots on the outside which show whether the site is significantly enriched after adjusting for multiple testing ($P < 3.59 \times 10^{-4}$). If there is a dot on the outermost ring then the site is significantly enriched when including all variants with $P < 5 \times 10^{-8}$ and if there is a dot on the inner ring then the site is enriched when including all variants with $P < 5 \times 10^{-5}$. If the site is significant for both thresholds there will be two dots.



AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

Figure E12 - Strength of association and model fit of the polygenic risk score in target dataset (UUS) by *P* threshold used

The x axis shows the *P* threshold (P_T) used for selecting variants to include in the risk score calculation. The black line and y axis on the left side shows the significance (−log(*P* value)) for the risk score for each P_T tested. The red dotted line shows the threshold of 0.001 used for determining whether the risk score was significantly associated with IPF susceptibility. The orange line and y axis on the right side shows the model fit (Nagelkerke's R^2) of the risk score at each P_T tested.



Supplementary References

1. Noth I, Zhang Y, Ma S, et al. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: A genome-wide association study. *The Lancet Respiratory Medicine*. 2013;1(4):309-317.

2. Fingerlin TE, Murphy E, Zhang W, et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet*. 2013;45(6):613-620.

3. Fingerlin TE, Zhang W, Yang IV, et al. Genome-wide imputation study identifies novel HLA locus for pulmonary fibrosis and potential role for auto-immunity in fibrotic idiopathic interstitial pneumonia. *BMC genetics*. 2016;17(1):1.

4. Allen RJ, Porte J, Braybrooke R, et al. Genetic variants associated with susceptibility to idiopathic pulmonary fibrosis in people of european ancestry: A genome-wide association study. *The Lancet Respiratory Medicine*. 2017.

5. Dressen A, Abbas AR, Cabanski C, et al. Analysis of protein-altering variants in telomerase genes and their association with MUC5B common variant status in patients with idiopathic pulmonary fibrosis: A candidate gene sequencing study. *The Lancet Respiratory Medicine*. 2018.

6. Wakefield J. A bayesian measure of the probability of false discovery in genetic epidemiology studies. *The American Journal of Human Genetics*. 2007;81(2):208-227.

7. Hao K, Bossé Y, Nickle DC, et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS Genet*. 2012;8(11):e1003029.

8. Lamontagne M, Couture C, Postma DS, et al. Refining susceptibility loci of chronic obstructive pulmonary disease with lung eqtls. *PLoS One*. 2013;8(7):e70220.

9. Obeidat M, Miller S, Probert K, et al. GSTCD and INTS12 regulation and expression in the human lung. *PLoS One*. 2013;8(9):e74630.

10. Jansen R, Hottenga J, Nivard MG, et al. Conditional eQTL analysis reveals allelic heterogeneity of gene expression. *Hum Mol Genet*. 2017;26(8):1444-1451.

11. GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature*. 2017;550(7675):204.

12. Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS genetics*. 2014;10(5):e1004383.

13. McLaren W, Gil L, Hunt SE, et al. The ensembl variant effect predictor. *Genome Biol*. 2016;17(1):122.

14. Ward LD, Kellis M. HaploReg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res*. 2011;40(D1):D930-D934.

AJRCCM Articles in Press. Published on 11-November-2019 as 10.1164/rccm.201905-1017OC

15. Zhou J, Troyanskaya OG. Predicting effects of noncoding variants with deep learning–based sequence model. *Nature methods*. 2015;12(10):931.

16. Dunham I, Kulesha E, Iotchkova V, Morganella S, Birney E. FORGE: A tool to discover cell specific enrichments of GWAS associated SNPs in regulatory regions. *F1000Research*. 2015;4.

17. Iotchkova V, Ritchie GR, Geihs M, et al. GARFIELD-GWAS analysis of regulatory or functional information enrichment with LD correction. *bioRxiv*. 2016:085738.

18. Slowikowski K, Hu X, Raychaudhuri S. SNPsea: An algorithm to identify cell types, tissues and pathways affected by risk loci. *Bioinformatics*. 2014;30(17):2496-2497.

19. Xu Y, Mizuno T, Sridharan A, et al. Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. *JCI insight*. 2016;1(20).

20. Shrine N, Guyatt AL, Erzurumluoglu AM, et al. New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nat Genet*. 2019:1.

21. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic risk score software. *Bioinformatics*. 2015;31(9):1466-1468.