Variants associated with {	extit{HHIP}} expression have sex-differential effects on lung function [version 1; peer review: 2 approved]

Katherine A. Fawcett\textsuperscript{1}, Ma'en Obeidat\textsuperscript{2}, Carl Melbourne\textsuperscript{1}, Nick Shrine\textsuperscript{1}, Anna L. Guyatt\textsuperscript{1}, Catherine John\textsuperscript{1}, Jian'an Luan\textsuperscript{3}, Anne Richmond\textsuperscript{4}, Marta R. Moksnes\textsuperscript{5}, Raquel Granell\textsuperscript{6}, Stefan Weiss\textsuperscript{7}, Medea Imboden\textsuperscript{8,9}, Sebastian May-Wilson\textsuperscript{10}, Pirro Hysi\textsuperscript{11}, Thibaud S. Boutin\textsuperscript{12}, Laura Portas\textsuperscript{12}, Claudia Flexeder\textsuperscript{13}, Sarah E. Harris\textsuperscript{14,15}, Carol A. Wang\textsuperscript{16}, Leo-Pekka Lyytikäinen\textsuperscript{17-19}, Teemu Palviainen\textsuperscript{20}, Rachel E. Foong\textsuperscript{21,22}, Dirk Keidel\textsuperscript{18,9}, Cosetta Minelli\textsuperscript{12}, Claudia Langenberg\textsuperscript{3}, Yohan Bossé\textsuperscript{12}, Maarten Van den Berge\textsuperscript{24}, Don D. Sin\textsuperscript{2,25}, Ke Hao\textsuperscript{26}, Archie Campbell\textsuperscript{17}, David Porteous\textsuperscript{27}, Sandosh Padmanabhan\textsuperscript{28}, Blair H. Smith\textsuperscript{19}, David M. Evans\textsuperscript{6,30,31}, Sue Ring\textsuperscript{6,30}, Arnulf Langhammer\textsuperscript{32}, Kristian Hveem\textsuperscript{5}, Cristen Willer\textsuperscript{33-35}, Ralf Ewert\textsuperscript{36}, Beate Stubbe\textsuperscript{36}, Nicola Pirastu\textsuperscript{10}, Lucija Klarić\textsuperscript{4}, Peter K. Joshi\textsuperscript{10}, Karina Patasova\textsuperscript{11}, Mangino Massimo\textsuperscript{11}, Ozren Polasek\textsuperscript{37}, John M. Starr\textsuperscript{14,38+}, Stefan Karrasch\textsuperscript{39-41}, Konstantin Strauch\textsuperscript{42,43}, Thomas Meitinginer\textsuperscript{44,45}, Igor Rudan\textsuperscript{10}, Taina Rantanen\textsuperscript{46}, Kirsi Pietiläinen\textsuperscript{47,48}, Mika Kähönen\textsuperscript{49,50}, Olli T. Raitakari\textsuperscript{51-53}, Graham L. Hall\textsuperscript{21,22}, Peter D. Sly\textsuperscript{54}, Craig E. Pennell\textsuperscript{16}, Jaakko Kaprio\textsuperscript{12,55}, Terho Lehtimäki\textsuperscript{17,18}, Veronique Vitart\textsuperscript{4}, Ian J. Deary\textsuperscript{14,15}, Debbie Jarvis\textsuperscript{12,56}, James F. Wilson\textsuperscript{4,10}, Tim Spector\textsuperscript{11}, Nicole Probst-Hensch\textsuperscript{8,9}, Nicholas J. Wareham\textsuperscript{3}, Henry Völzke\textsuperscript{57}, John Henderson\textsuperscript{30+}, David P. Strachan\textsuperscript{58}, Ben M. Brumpton\textsuperscript{5,59,60}, Caroline Hayward\textsuperscript{16}, Ian P. Hall\textsuperscript{61}, Martin D. Tobin\textsuperscript{1,62}, Louise V. Wain\textsuperscript{1,62}

\textsuperscript{1}Department of Health Sciences, University of Leicester, Leicester, LE1 7RH, UK
\textsuperscript{2}The University of British Columbia Centre for Heart Lung Innovation, St Paul's Hospital, Vancouver, BC, Canada
\textsuperscript{3}MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, CB2 0QQ, UK
\textsuperscript{4}MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh, EH4 2XU, UK
\textsuperscript{5}K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway
\textsuperscript{6}Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN, UK
\textsuperscript{7}Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, 17475, Germany
\textsuperscript{8}Swiss Tropical and Public Health Institute, Basel, Switzerland
\textsuperscript{9}University of Basel, Basel, Switzerland
\textsuperscript{10}Centre for Global Health Research, Usher Institute, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, UK
The Department of Twin Research & Genetic Epidemiology, King's College London, St Thomas' Campus, Lambeth Palace Road, London, UK

Population Health and Occupational Disease, National Heart and Lung Institute, Imperial College London, London, UK

Institute of Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, 85764, Germany

Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, EH8 9JZ, UK

Psychology, University of Edinburgh, Edinburgh, EH8 9JZ, UK

School of Medicine and Public Health, Faculty of Medicine and Health, The University of Newcastle, Callaghan, Australia

Department of Clinical Chemistry, Fimlab Laboratories, Tampere, 33520, Finland

Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, 33014, Finland

Department of Cardiology, Heart Center, Tampere University Hospital, Tampere, 33521, Finland

Institute for Molecular Medicine FIMM, University of Helsinki, Helsinki, FI-00014, Finland

Telethon Kids Institute, Perth, Australia

School of Physiotherapy and Exercise Science, Faculty of Health Sciences, Curtin University, Perth, Australia

Institut universitaire de cardiologie et de pneumologie de Québec, Department of Molecular Medicine, Laval University, Québec, Canada

University Medical Center Groningen, Department of Pulmonology, GRIAC Research Institute, University of Groningen, Groningen, The Netherlands

Respiratory Division, Department of Medicine, University of British Columbia, Vancouver, BC, Canada

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

Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, EH4 2XU, UK

British Heart Foundation Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, G12 8TA, UK

Division of Population Health Sciences, Ninewells Hospital and Medical School, University of Dundee, Dundee, DD1 9SY, UK

Population Health Sciences Bristol Medical School, University of Bristol, Bristol, BS8 2BN, UK

University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, QLD 4072, Australia

Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology, Trondheim, Norway

Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, USA

Department of Internal Medicine, University of Michigan, Ann Arbor, USA

Department of Human Genetics, University of Michigan, Ann Arbor, USA

Department of Internal Medicine B, Cardiology, Pneumology, Infectious Diseases, Intensive Care Medicine, University Medicine Greifswald, Greifswald, 17475, Germany

University of Split School of Medicine, Split, Croatia

Alzheimer Scotland Research Centre, University of Edinburgh, Edinburgh, EH8 9JZ, UK

Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, Ludwig-Maximilians-Universität, Munich, 80336, Germany

Institute of Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, 85764, Germany

Comprehensive Pneumology Center Munich (CPC-M), Member of the German Center for Lung Research (DZL), Munich, 81377, Germany

Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany

Chair of Genetic Epidemiology, IBE, Faculty of Medicine, LMU Munich, Munich, 81377, Germany

Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany

Institute of Human Genetics, Klinikum rechts der Isar der TU Muenchen, Muenchen, 81675, Germany

Faculty of Sport and Health Sciences, Gerontology Research Center, University of Jyväskylä, Jyväskylä, Finland

Obesity Research Unit, Research Program for Clinical and Molecular Metabolism, Faculty of Medicine, University of Helsinki, Helsinki, FI-00014, Finland

Obesity Centre, Abdominal Centre, Helsinki University Hospital and University of Helsinki, Helsinki, FI-00029, Finland

Department of Clinical Physiology, Tampere University Hospital, Tampere, 33521, Finland

Department of Clinical Physiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, 33014, Finland
Abstract

**Background:** Lung function is highly heritable and differs between the sexes throughout life. However, little is known about sex-differential genetic effects on lung function. We aimed to conduct the first genome-wide genotype-by-sex interaction study on lung function to identify genetic effects that differ between males and females.

**Methods:** We tested for interactions between 7,745,864 variants and sex on spirometry-based measures of lung function in UK Biobank (N=303,612), and sought replication in 75,696 independent individuals from the SpiroMeta consortium.

**Results:** Five independent single-nucleotide polymorphisms (SNPs) showed genome-wide significant (P<5x10^{-8}) interactions with sex on lung function, and 21 showed suggestive interactions (P<1x10^{-6}). The strongest signal, from rs7697189 (chr4:145436894) on forced expiratory volume in 1 second (FEV\textsubscript{1}) (P=3.15x10^{-15}), was replicated (P=0.016) in SpiroMeta. The C allele increased FEV\textsubscript{1} more in males (untransformed FEV\textsubscript{1} β=0.028 [SE 0.0022] litres) than females (β=0.009 [SE 0.0014] litres), and this effect was not accounted for by differential effects on height, smoking or pubertal age. rs7697189 resides upstream of the hedgehog-interacting protein (HHIP) gene and was previously associated with lung function and HHIP lung expression. We found HHIP expression was significantly different between the sexes (P=6.90x10^{-6}), but we could not detect sex differential effects of rs7697189 on expression.

**Conclusions:** We identified a novel genotype-by-sex interaction at a putative enhancer region upstream of the HHIP gene. Establishing the mechanism by which HHIP SNPs have different effects on lung function in males and females will be important for our understanding of lung health and diseases in both sexes.

**Keywords**

genome-wide interaction study, lung function, sex, HHIP, expression
Corresponding authors: Katherine A. Fawcett (kaf19@leicester.ac.uk), Louise V. Wain (lwvl@leicester.ac.uk)


Competing interests: The following authors report potential competing interests: L.V.W.: Louise V. Wain has received grant support from GSK. M.D.T.: Martin D. Tobin has received grant support from GSK. I.P.H.: Ian P. Hall has received support from GSK and BI.

Grant information: This work was supported by the Wellcome Trust through a Wellcome Trust/BHF Fellowship awarded to A.L.G.; an Investigator Award awarded to M.D.T. (202849); core support for ALSPAC (102215); a Strategic Award as support for Generation Scotland (104036); and GABRIEL project funding as support for SAPALDIA (084703). Sources of support: K.A.F. holds an Asthma UK fellowship. A.L.G. is supported by a Wellcome Trust/BHF fellowship. C.J. holds a Medical Research Council (MRC) Clinical Research Training Fellowship (MR/P00167X/1). L.V.W. holds a GSK/British Lung Foundation Chair in Respiratory Research. M.D.T. is supported by a Wellcome Trust Investigator Award (202849). M.D.T. and L.V. Wain have been supported by the MRC (MR/N011317/1). The research was partially supported by the NIHR Leicester 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.P.H.: The research 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. Study-specific sources of support: ALSPAC: The UK Medical Research Council and Wellcome (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. GWAS data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf); Lung function measurements at 24 years were specifically funded by MRC MR/M022501/1. This publication is

Page 4 of 19
Introduction

Measures of lung function, including forced expiratory volume in 1 second (FEV₁) and forced vital capacity (FVC), are used to determine diagnosis and severity of chronic obstructive pulmonary disease (COPD). COPD refers to a group of complex lung disorders characterised by irreversible (and usually progressive) airway obstruction, and is projected to be the third leading cause of death globally in 2020. The major risk factor for COPD is smoking, but other environmental and genetic factors have been identified.

Physiological lung development and function differ throughout life between males and females. It is known that sex hormones can influence these processes but the mechanisms are not well understood. The incidence and presentation of lung diseases such as COPD also exhibit sexual dimorphism. Traditionally viewed as a disease of older males, COPD has been increasing in prevalence amongst females over the last two decades. It has been reported that females are more vulnerable to environmental risk factors for COPD and are over-represented amongst sufferers of early-onset severe COPD. Females are also more likely to present with small airway disease whereas males are more likely to develop emphysematous phenotype. Moreover, females report more frequent and/or severe exacerbations of respiratory symptoms than males and higher levels of dyspnoea and cough.

In a recent paper, 279 genetic loci were reported as associated with lung function traits, but these only explain a small proportion of the heritability. One possible source of hidden heritability is the interaction between genetic factors and biological sex on lung function traits. A genome-wide genotype-by-sex interaction study in three studies comprising 6260 COPD cases and 5269 smoking controls found a putative sex-specific risk factor for COPD in the CELSR1 gene, a region not previously implicated in COPD or lung function. However, having sufficient statistical power to reproducibly detect genotype-by-sex interactions requires much larger sample sizes. Statistical power can also be enhanced by using quantitative lung function traits as outcomes instead of COPD diagnoses, but we are not aware of any genome-wide genotype-by-sex interaction studies on lung function traits. Understanding the role of sex in lung function and COPD will be important for developing therapeutics that work for both males and females.

In this study, we tested for an interaction effect of 7,745,864 variants and sex on FEV₁, FEV₁/FVC, FVC, and PEF in 303,612 individuals from the UK Biobank resource. We sought replication of our findings in 75,696 independent individuals from the SpiroMeta consortium. To our knowledge this is the first genome-wide sex-by-genotype interaction study on lung function traits, and the largest sex-by-genotype interaction study to focus on COPD-related outcomes.

Results

We tested 7,745,864 genome-wide variants with minor allele frequency (MAF) ≥ 0.01 and imputation quality scores ≥ 0.3 for genotype-by-sex interactions on lung function in 303,612 unrelated individuals of European ancestry from UK Biobank. Five independent signals were identified showing genome-wide significant (P<5 x 10⁻⁸) interaction with sex on at least one of four lung function traits (FEV₁, FEV₁/FVC, FVC, and PEF) with a further 21 SNPs showing suggestive significance (P<1 x 10⁻⁴) (Table 1; Figure S1, Extended data). The top three genome-wide significant signals had been previously reported for association with lung function: rs7697189 near the gene encoding hedgehog-interacting protein (HHIP) (interaction P = 3.15 x 10⁻¹⁰), rs9403386 near the gene encoding Adhesion G Protein-Coupled Receptor G6 (ADGRG6, previously known as GPR126) (interaction P = 4.56 x 10⁻⁸), and rs162185 downstream of the gene encoding transcription factor 21 (TCF21) (interaction P = 4.87 x 10⁻⁷). This may, in part, reflect greater power to detect interactions with variants with strong main effects on lung function. Only rs355079 (interaction P = 8.84 x 10⁻⁷) showed significant effects in opposite directions in males compared to females.

We sought evidence for replication of all 26 signals in up to 75,696 individuals from 20 cohorts of the SpiroMeta consortium. One variant, rs76911399, was excluded because it was poorly imputed in SpiroMeta cohorts and had no directly genotyped or well-imputed proxies (at r² threshold 0.8). Of the remaining 25 signals, 19 exhibited the same direction of interaction effect as in UK Biobank. Furthermore, the effect sizes (beta coefficients) from the regression analyses of all 25 SNPs in UK Biobank and SpiroMeta showed a correlation of 0.51 (Figure S2, Extended data). The SNP with the strongest evidence for interaction with sex on lung function in SpiroMeta cohorts was rs7697189 (near HHIP) (replication interaction P = 0.016) (Table 1, Figure 1). The minor (C) allele of rs7697189 had a larger effect on lung function in males (β = 0.052 [SE 0.004], P = 2.13 x 10⁻³) compared to females (β = 0.013 [SE 0.003], P = 1.16 x 10⁻³) (Table 1). This SNP resides upstream of the HHIP gene and is in linkage disequilibrium with two previously reported lung function-associated sentinel SNPs, rs13141644 (r² = 0.91) and rs13116999 (r² = 0.56). SNP rs7697189 also showed some evidence of interaction with sex on PEF (β = -0.035 (0.005), P = 8.78 x 10⁻⁵), FEV₁/FVC (β = -0.028 (0.005), P = 8.98 x 10⁻⁴), and FVC (β = -0.020 (0.005), P = 8.71 x 10⁻⁴) (Table S1, Extended data).

rs7697189 interacts with sex on lung function independently of height, smoking and pubertal timing

As SNPs in HHIP are also reported to be associated with height and increased height is associated with increased lung function, it is possible that rs7697189 has differential effects on lung function in males and females through differential effects on height. However, the association of rs7697189 with standing height was not modified by sex in a combined analysis of UK Biobank males and females with a genotype-by-sex interaction term (interaction P = 0.806). We also conducted a sensitivity analysis showing that the effect of the rs7697189-by-sex interaction on FEV₁ was consistent with the original estimate after adjustment for sitting height (β = -0.04 [SE = 0.005], P = 1.97 x 10⁻¹⁵).

Amongst the 303,612 UK Biobank participants in this study, the proportion of ever-smokers was higher in males (52.8%) than females (40.3%) (Table S2). A larger effect of rs7697189 on
Table 1. Association between top SNPs and lung function in males and females, and genotype-by-sex interaction results.

<table>
<thead>
<tr>
<th>SNP (nearest gene and coordinates)</th>
<th>Test/other allele</th>
<th>Trait</th>
<th>Lung function UK Biobank males</th>
<th>Lung function UK Biobank females</th>
<th>Sex interaction in UK Biobank</th>
<th>Sex interaction in SpiroMeta</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>MAF</td>
<td>Beta (SE)</td>
<td>P</td>
<td>MAF</td>
</tr>
<tr>
<td>rs7697189 (HHIP) 4:145436994</td>
<td>C/G</td>
<td>FEV₁</td>
<td>0.390</td>
<td>0.052 (0.004)</td>
<td>2.13E-33</td>
<td>0.392</td>
</tr>
<tr>
<td>rs9403836 (ADGRG6) 6:142764073</td>
<td>C/A</td>
<td>FEV₁/FVC</td>
<td>0.031</td>
<td>0.214 (0.012)</td>
<td>4.48E-75</td>
<td>0.031</td>
</tr>
<tr>
<td>rs162185 (TCF21) 6:134226147</td>
<td>C/T</td>
<td>PEF</td>
<td>0.411</td>
<td>-0.038 (0.004)</td>
<td>1.35E-18</td>
<td>0.410</td>
</tr>
<tr>
<td>rs6480592 (CHST3) 10:73764509</td>
<td>C/T</td>
<td>PEF</td>
<td>0.398</td>
<td>-0.021 (0.004)</td>
<td>1.66E-06</td>
<td>0.400</td>
</tr>
<tr>
<td>rs111893604 (ZSCAN10) 16:3141104</td>
<td>G/T</td>
<td>FEV₁</td>
<td>0.059</td>
<td>0.040 (0.009)</td>
<td>1.70E-05</td>
<td>0.059</td>
</tr>
<tr>
<td>rs72694266 (RP11-907D1.1) 14:97578576</td>
<td>A/C</td>
<td>PEF</td>
<td>0.077</td>
<td>-0.044 (0.008)</td>
<td>2.69E-07</td>
<td>0.078</td>
</tr>
<tr>
<td>rs72781459 10:10247676</td>
<td>C/T</td>
<td>PEF</td>
<td>0.096</td>
<td>0.031 (0.007)</td>
<td>3.44E-05</td>
<td>0.097</td>
</tr>
<tr>
<td>rs74316059 (RP11-649A16.1) 3:146983325</td>
<td>T/C</td>
<td>FEV₁/FVC</td>
<td>0.042</td>
<td>0.049 (0.010)</td>
<td>2.52E-06</td>
<td>0.043</td>
</tr>
<tr>
<td>rs55789572 (EIF2S2/RALY) 20:32687822</td>
<td>A/C</td>
<td>FEV₁</td>
<td>0.022</td>
<td>0.041 (0.015)</td>
<td>0.006</td>
<td>0.022</td>
</tr>
<tr>
<td>rs74935318 (DAPK2) 15:64303295</td>
<td>A/G</td>
<td>PEF</td>
<td>0.025</td>
<td>-0.072 (0.014)</td>
<td>1.23E-07</td>
<td>0.025</td>
</tr>
<tr>
<td>rs11247571 (ABR) 17:908502</td>
<td>G/A</td>
<td>PEF</td>
<td>0.343</td>
<td>-0.025 (0.005)</td>
<td>3.65E-08</td>
<td>0.344</td>
</tr>
<tr>
<td>rs707588 (RP11-154H17.1) 1:57111430</td>
<td>G/A</td>
<td>FEV₁</td>
<td>0.482</td>
<td>-0.020 (0.004)</td>
<td>3.23E-06</td>
<td>0.482</td>
</tr>
<tr>
<td>rs138473298 (AUTS2) 7:69644989</td>
<td>T/C</td>
<td>PEF</td>
<td>0.012</td>
<td>-0.077 (0.020)</td>
<td>0.0002</td>
<td>0.011</td>
</tr>
<tr>
<td>rs139069254 (RP11-648K4.2) 15:88113916</td>
<td>A/G</td>
<td>FEV₁</td>
<td>0.018</td>
<td>0.071 (0.016)</td>
<td>1.83E-05</td>
<td>0.018</td>
</tr>
<tr>
<td>SNP (nearest gene) and coordinates</td>
<td>Test/other allele</td>
<td>Trait</td>
<td>MAF</td>
<td>Beta (SE)</td>
<td>P</td>
<td>MAF</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-----------------</td>
<td>------</td>
<td>-----</td>
<td>----------</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>rs138163836 (PVRL3) 3:110952902</td>
<td>C/T</td>
<td>FVC</td>
<td>0.021</td>
<td>0.064 (0.015)</td>
<td>1.94E-05</td>
<td>0.020</td>
</tr>
<tr>
<td>rs2843055 (XDH) 2:31573390</td>
<td>T/G</td>
<td>FEV&lt;sub&gt;1&lt;/sub&gt;</td>
<td>0.012</td>
<td>0.065 (0.020)</td>
<td>0.002</td>
<td>0.013</td>
</tr>
<tr>
<td>rs117380804 18:76145905</td>
<td>T/C</td>
<td>FVC</td>
<td>0.035</td>
<td>0.035 (0.012)</td>
<td>0.003</td>
<td>0.036</td>
</tr>
<tr>
<td>rs602622 (RASGRP3) 2:3368226</td>
<td>C/G</td>
<td>PEF</td>
<td>0.444</td>
<td>-0.022 (0.004)</td>
<td>2.11E-07</td>
<td>0.445</td>
</tr>
<tr>
<td>rs2253718 (RF00019, SFTA2) 6:30900427</td>
<td>T/G</td>
<td>PEF</td>
<td>0.409</td>
<td>-0.049 (0.004)</td>
<td>5.69E-30</td>
<td>0.405</td>
</tr>
<tr>
<td>rs2353939 (HHIP) 4:145729724</td>
<td>G/A</td>
<td>FVC</td>
<td>0.437</td>
<td>0.016 (0.004)</td>
<td>0.002</td>
<td>0.435</td>
</tr>
<tr>
<td>rs7691139 (ZNF280A) 22:220976151</td>
<td>G/C</td>
<td>FEV&lt;sub&gt;1&lt;/sub&gt;/FVC</td>
<td>0.116</td>
<td>-0.025 (0.007)</td>
<td>0.003</td>
<td>0.115</td>
</tr>
<tr>
<td>rs10003954 2:17296894</td>
<td>C/T</td>
<td>FVC</td>
<td>0.014</td>
<td>0.050 (0.017)</td>
<td>0.004</td>
<td>0.014</td>
</tr>
<tr>
<td>rs7371120 (MLF1) 3:156097633</td>
<td>A/C</td>
<td>FVC</td>
<td>0.346</td>
<td>0.029 (0.004)</td>
<td>3.72E-11</td>
<td>0.346</td>
</tr>
<tr>
<td>rs355079 (LMCD1-AS1) 3:8643371</td>
<td>T/C</td>
<td>FVC</td>
<td>0.337</td>
<td>0.015 (0.004)</td>
<td>0.0007</td>
<td>0.339</td>
</tr>
<tr>
<td>rs7338055 (SPRYD7) 13:50504226</td>
<td>T/C</td>
<td>FVC</td>
<td>0.259</td>
<td>0.018 (0.005)</td>
<td>0.0001</td>
<td>0.259</td>
</tr>
<tr>
<td>rs34490170 (NEUROD1/ CERKL) 2:182576419</td>
<td>C/T</td>
<td>FVC</td>
<td>0.110</td>
<td>-0.035 (0.007)</td>
<td>6.41E-07</td>
<td>0.110</td>
</tr>
</tbody>
</table>

The SNPs are those that demonstrate a sex-interaction effect on lung function in UK Biobank (P < 1x10^{-6}) (N = 303,612). Lung function traits were pre-adjusted for age, age², standing height and smoking status and the residuals rank-transformed to normality. The regression models also included genotyping array and the first ten ancestry-based principal components. For each SNP, columns 4-9 provide minor allele frequency (MAF), and beta-coefficients, standard errors and the P value for their association with lung function in males and females separately. Columns 10-11 show the results of the SNP-by-sex interaction in UK Biobank, where the effect is given in females relative to males. For example, the top SNP (rs7691139) shows a less positive effect in females compared to males and its beta coefficient is therefore negative. Columns 12-13 show the results of the SNP-by-sex interaction in 20 cohorts of the SpiroMeta consortium (N = 75,696). Bold text in the final column indicates the effect in SpiroMeta was in the same direction to the effect in UK Biobank.
Figure 1. Meta-analysis of rs7697189-by-sex interaction effects on lung function in SpiroMeta cohorts. The forest plot shows the beta-coefficients (test effects, TE) and standard errors for the interaction between rs7697189 and sex on forced expiratory volume in 1 second (FEV$_1$) in 20 cohorts of the SpiroMeta consortium (total N = 75,696). The overall effect size from fixed effects meta-analysis is represented by the diamond.

SNP rs7697189, and correlated SNPs in the region, have been shown to be associated with expression levels of HHIP in lung tissue$^{19}$.

HHIP is a critical protein during early development and HHIP variants have been associated with lung function in infancy$^{20}$. We tested whether HHIP SNPs also have differential effects on lung function in females compared to males in childhood using data from children with an average age of eight years in the ALSPAC and Raine studies (N = 5645). In the meta-analysis of ALSPAC and Raine (Figure S3, Extended data$^{10}$), whilst we observed a point estimate for the rs7697189-by-sex interaction effect on FEV$_1$ that was consistent with the confidence intervals for the discovery effect observed in UK Biobank, the confidence intervals overlapped the null (which likely reflects in part the smaller numbers studied in these cohorts). Finally, as pubertal timing has been associated with adult lung function$^{21}$, we tested for an effect of relative age at puberty on the association between rs7697189 and lung function in a sex-stratified analysis. The association between HHIP SNPs and lung function was adjusted for relative age at voice breaking in males and for age at menarche in females, but adjusted effect estimates were highly consistent with the unadjusted estimates of the SNPs on lung function (Table S3, Extended data$^{10}$).

rs7697189 is associated with HHIP expression, but no interaction with sex

It is possible that rs7697189 interacts with sex on lung function through differential effects on HHIP expression. We confirmed that rs7697189 is associated with HHIP expression in lung tissue
Figure 2. Genotype-by-sex interaction results within the HHIP region for lung function traits in UK Biobank. The SNP with the strongest association in the rs7697189-proximal region is represented by a blue diamond. The FEV₁ and PEF sentinels are rs7697189, the FEV₁/FVC sentinel is rs1512281 ($R^2 = 0.95$ with rs7697189), and the FVC sentinel is rs7681384 ($R^2 = 0.57$ with rs7697189). Note that there is an independent suggestively significant signal from rs2353939 and surrounding SNPs for FVC, but this did not replicate in SpiroMeta cohorts. All other SNVs are colour coded according to their linkage disequilibrium ($R^2$) with the sentinel SNP (as shown in the key). All imputed SNVs are plotted irrespective of MAF, demonstrating that rarer variants are not exhibiting significant interactions with sex on lung function. The locations of genes in the region are shown in the lower panel of each plot. Recombination rate is represented by the blue lines. These plots were generated using LocusZoom software.
but we did not detect an interaction with sex on HHIP expression (Table S4, Extended data)\(^{19}\). However, HHIP (in all samples irrespective of genotype at rs7697189) did show differential expression between males and females, with females showing higher expression (Table S5; Extended data)\(^{18}\). This agrees with GTEx data on HHIP lung expression in males and females (Figure S4, Extended data)\(^{18}\).

rs7697189 is in linkage disequilibrium with a SNP predicted to disrupt SREBP and SRF motifs

HaploReg v4.1\(^{22}\) was used to identify whether rs7697189, or SNPs in linkage disequilibrium, affected transcription factor binding motifs. This demonstrated that rs7697189 itself was predicted to change FAC1 and FOXO motifs and was within a chromatin mark indicative of enhancer activity in embryonic stem cell lines differentiated to CD56+ mesoderm and CD184+ endoderm cultured cells. A SNP (rs12504628) in complete linkage disequilibrium with rs7697189 changes SREBP and SRF motifs. These transcription factors have been reported to be involved in sex hormone signalling\(^{23,24}\).

### Discussion

We identified a genome-wide significant genotype-by-sex interaction signal at a locus previously reported for association with lung function upstream of the HHIP gene (rs7697189, FEV\(_1\) interaction \(P = 3.15 \times 10^{-13}\)). The SNP showed some evidence of replication in 75,696 individuals from 20 independent studies of the SpiroMeta consortium (\(\beta = -0.025\) (0.01), \(P = 0.016\)), although it did not pass a Bonferroni correction for multiple testing. We demonstrated that the differential effects of this SNP in males and females (FEV\(_1\), \(\beta = 0.052\) (0.004) in males and 0.013 (0.003) in females, corresponding to an untransformed FEV\(_1\), \(\beta = 0.028\) [SE 0.0022] litres in males vs \(\beta = 0.009\) [SE 0.0014] litres in females) did not appear to be mediated by effects on height, smoking behaviour or pubertal age.

There was evidence that SNPs at the HHIP locus demonstrated interactions with sex on two additional lung function traits in UK Biobank: FEV\(_1\)/FVC and PEF (\(\beta = -0.028\) (0.005), \(P = 8.78 \times 10^{-12}\) and \(\beta = -0.035\) (0.005), \(P = 8.78 \times 10^{-12}\), respectively). Stratified analyses in males and females demonstrated that these SNPs appeared to have a stronger effect on lung function in males compared to females. There was no interaction between these SNPs and ever-smoking status on lung function in UK Biobank, suggesting that the stronger effect in males is not due to differences in smoking behaviour. We also demonstrate that an association between these SNPs and height is not modified by sex, suggesting that differential effects on height in males and females do not explain the genotype-by-sex interaction on lung function.

In contrast to these results, a recent study found comparatively weak evidence of an interaction effect between a SNP (rs13140176) in high LD with rs7697189 (\(r^2 = 0.93\)) and sex on risk of COPD in UK Biobank\(^{25}\). This is likely in part to be due to reduced power to detect interaction effects on a binary trait. Indeed, in our study, the rs13140176-by-sex interaction effect on FEV\(_1\)/FVC passes the conventional threshold for genome-wide significance (\(P<5\times10^{-8}\)) but when COPD was defined as FEV\(_1\)/FVC<0.7 this threshold was not met (\(P=0.023\)). Nevertheless, rs13140176 shows a consistent direction of effect between the studies: the lung function-lowering allele increases risk of COPD to a greater extent in males than females\(^{26}\).

The genome-wide significant sex interaction locus is located upstream of the HHIP gene, a region previously reported to be associated with lung function\(^{12,15}\) and HHIP gene expression\(^{19}\). The HHIP gene encodes hedgehog-interacting protein, a negative regulator of hedgehog signalling. The hedgehog signalling pathway regulates numerous physiological processes such as growth, self-renewal, cell survival, differentiation, migration, and tissue polarity and plays a vital role in the morphogenesis of lung and other organs\(^{26}\). Hedgehog signalling has also been shown to participate in regulation of stem and progenitor cell populations in adult tissues, impacting tissue homeostasis and repair\(^{27}\). SNP rs7697189, showing the strongest sex interaction on lung function in our study, is in strong linkage disequilibrium (\(R^2 = 0.93\)) with SNPs residing in an HHIP enhancer region\(^{28}\). These enhancer-region SNPs were reported to be associated with enhancer activity and HHIP expression in lung tissues. They also exhibit genome-wide significant genotype-by-sex interactions on lung function in our data. We therefore tested the effect of rs7697189 on HHIP expression in lung tissue from 472 males and 566 females to look for sex differential effects. In contrast to the previous study\(^{28}\), we found that the lung-function lowering G allele was associated with enhanced expression of HHIP in both males and females, and that expression was lower in males than females. However, the association between rs7697189 and HHIP expression was not modified by sex. This may be because there is no sex differential effect on expression, or the study might have been underpowered to detect an interaction effect. It is therefore still not clear why SNPs upstream of HHIP would be showing different effects in males and females. Our \textit{in silico} analyses predict that rs7697189 and a SNP in linkage disequilibrium (rs12504628) change transcription factor motifs that may be relevant to the effect of sex hormones on lung development, but experimental analyses will be required to test these hypotheses.

Investigating the effects of HHIP at different stages of development by sex may help to shed light on its mechanism of action. In our study we had access to genetic and lung function data from 5645 children with an average age of eight years. Though underpowered to detect the association between rs7697189 and FEV\(_1\) seen in UK Biobank adults, the lack of a similar trend in children suggests that HHIP variants may have differential effects at different developmental stages (though the genotype-by-sex interaction is in the same direction as in adults). We also looked for an effect of timing of puberty on the association between rs7697189 and lung function in adults, but adjustment for relative age of voice breaking in males and relative age at menarche in females made no difference to the relationship between rs7697189 and lung function. As UK Biobank participants were aged between 40 and 69 years at recruitment, we did not have the longitudinal data to investigate the effect of HHIP SNPs on trajectories of lung function decline throughout life\(^{28}\), but this could be an interesting area for future studies.

We identified four additional genome-wide significant (interaction \(P<5\times10^{-8}\)) sex-by-genotype interactions on lung function in our discovery analysis in UK Biobank, with a further 21 that met a less stringent threshold of interaction (\(P<1\times10^{-6}\)). As far as we are
aware, this is the first genome-wide sex-by-genotype interaction study for lung function traits. We did not find a significant genotype-by-sex interaction on lung function or COPD at the CELSR1 locus (interaction P = 0.525 and P = 0.503, respectively) previously reported to have sex-specific effects on risk of COPD.

In conclusion, we have identified a novel genotype-by-sex interaction at SNPs at a putative enhancer region upstream of the hedgehog-interacting protein (HHIP) gene. Establishing the mechanism by which HHIP has sex differential effects on lung function will be important for our understanding of the biological underpinnings of COPD in males and females. This knowledge, in turn, will be crucial to optimising treatment in males and females.

Materials and Methods

Ethics and consent
This study used anonymised data from UK Biobank (RRID: SCR_012815), which comprises over 500,000 volunteer participants aged 40–69 years recruited across Great Britain between 2006 and 2010. The protocol and consent were approved by the UK Biobank’s Research Ethics Committee. Our analysis was conducted under approved UK Biobank data application number 648. For SpiroMeta consortium cohorts, all participants provided written informed consent and studies were approved by local Research Ethics Committees and/or Institutional Review boards. Full ethics statements for each SpiroMeta consortium cohort is included in the S1 Appendix (Extended data).

UK Biobank
The UK Biobank is described here: http://www.ukbiobank.ac.uk. Individuals were included in this study if (i) they had no missing data for sex, age, height, and smoking status, (ii) their spirometry data passed quality control, as described previously, (iii) their genetically inferred sex matched their reported sex, (iv) they had genome-wide imputed genetic data, (v) they were of genetically determined European ancestry, and (vi) they were not first- or second-degree relatives of any other individual included in the study. In total, 303,612 individuals met these criteria (Table S2, Extended data).

Participants’ DNA was genotyped using either the Affymetrix Axiom® UK BiLEVE array or the Affymetrix Axiom® UK Biobank array. Genotypes were imputed based on the Human Reference Consortium (HRC) panel, as described elsewhere. Variants with minor allele frequency (MAF)<0.01 were excluded, as were variants with imputation quality scores <0.3.

SpiroMeta consortium
The SpiroMeta consortium meta-analysis comprised 75,696 individuals from 20 studies (see S1 Appendix for details, Extended data). Ten studies (N=17,280) were imputed using 1000 Genomes Phase 1 reference panel, nine (N=37,919) were imputed using the HaploType Reference Consortium (HRC) panel, and one (N=2077) was imputed using the HapMap CEU Build 36 Release 22. The ALSPAC (RRID: SCR_007260) and Raine studies also provided data on children with an average age of eight years (N=4426 and N=1219, respectively). Tables S6 and S7 show definitions of all abbreviations, study characteristics, details of genotyping platforms and imputation panels and methods (Extended data). Measurements of spirometry for each study are as previously described. Fourteen SpiroMeta studies had data on PEF (N=51,555).

Statistical analysis

Spirometry-based lung function traits FEV1, FEV1/FVC, FVC, and PEF were pre-adjusted for age, age2, standing height (or sitting height in the sensitivity analysis) and smoking status and the residuals rank-transformed to normality using the rntrans function of the GenABEL package (RRID: SCR_001842) in R (RRID: SCR_001905). To test each imputed autosomal variant for an interaction effect, a linear regression model with genotype (additive effect), sex, genotype-by-sex interaction, genotyping array and the first ten principal components included as covariates was implemented using Plink 2.0 software (RRID: SCR_001757). Step-wise conditional analyses to identify independently associated variants were undertaken using GCTA software.

Regression analysis to test genotype-by-sex interactions on height were conducted using a model including genotype (additive effect), age, age2, sex, genotyping array and the first ten principal components as covariates. Interactions between smoking status and genotype on lung function were tested using lung function traits transformed as described above (with sex included in the model instead of ever-smoking status). The linear regression model included genotype (additive effect), ever-smoking status, a genotype-by-smoking interaction term, genotyping array and the first ten principal components.

To test whether pubertal timing has differential effects on the association between SNPs and lung function in males and females, the regression model was adjusted for relative age at menarche in females and relative age at voice breaking in males. Relative age at voice breaking is categorised as earlier than average (1), around average (2) and later than average (3) in UK Biobank. Age at menarche is given as the participant’s age at menarche in years. To make these variables comparable, age at menarche was categorised as early (<12 years old), average (12–14 years old) and late (>14 years old) as in a previous study. As in the lung function analyses, ancestry-based principal components and genotyping array were included in all the regression models.

For the SpiroMeta consortium, summary statistics were generated by each contributing cohort separately according to the same analysis plan as the UK Biobank data. Meta-analysis of SpiroMeta cohorts was conducted using inverse-variance weighted fixed effects meta-analysis using the metagen function of the meta package in R.

The lung eQTL study

The lung expression quantitative trait loci (eQTL) study database has been described previously and in S1 Appendix (Extended data). HHIP differential gene expression analysis between females and males was performed using linear regression. Association of rs7697189 and rs7697189-by-sex interaction with gene expression was tested in 1,038 subjects with genotypes
using MatrixEQTL package in R. All analyses were done separately in Laval, UBC and Groningen, and then combined using a meta-analysis with fixed-effects model and inverse-variance weights.

**Data availability**

**Underlying data**

UK Biobank data is an open access resource available to bona fide researchers undertaking health-related research. Researchers must apply for access (see [https://www.ukbiobank.ac.uk/researchers](https://www.ukbiobank.ac.uk/researchers) for more details). Genome-wide interaction study summary statistics are available on Figshare (see below).

Figshare: Genome-wide sex interaction study summary statistics for lung function traits in UK Biobank. [https://doi.org/10.6084/m9.figshare.12298736.v1](https://doi.org/10.6084/m9.figshare.12298736.v1)

Extended data

Figshare: Variants associated with HHIP expression have sex-differential effects on lung function: supplementary material. [https://doi.org/10.6084/m9.figshare.1212920](https://doi.org/10.6084/m9.figshare.1212920)

This project contains Fawcett_et_al_Extended_data_supplement.docx, which contains the following extended data:

- Supplementary materials and methods
- Figure S1. Genome-wide interaction SNP-by-sex interaction results on four measures of lung function in UK Biobank
- Figure S2. Correlation between genotype-by-sex interaction effect sizes in UK Biobank and the SpiroMeta studies
- Figure S3. Association between rs7697189 and FEV₁ in children from the ALSPAC and Raine cohorts
- Figure S4. GTEx data on expression of HHIP by sex in different tissues
- Table S1. Association between rs7697189 and lung function traits in males and females, and genotype-by-sex interaction results
- Table S2. UK Biobank demographics
- Table S3. Sex-stratified association between rs7697189 and lung function before and after adjustment for pubertal timing
- Table S4. Association between rs7697189 and HHIP expression and rs7697189-by-sex interaction on HHIP expression
- Table S5. Differential expression of HHIP in males compared to females
- Table S6. SpiroMeta studies
- Table S7. SpiroMeta analysis methods

Data are available under the terms of the Creative Commons Zero “No rights reserved” data waiver (CC0 1.0 Public domain dedication).

**Acknowledgements**

We gratefully acknowledge the contributions of co-authors Professor John M. Starr and Professor John Henderson, both of whom died prior to the publication of this manuscript. We thank UK Biobank and all the participants for generating this important health research resource. This study used the ALICE and SPEC-TRE High Performance Computing Facilities at the University of Leicester. The ALSPAC study team are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The ECRHS study would like to thank the participants, field workers and researchers who have participated in the ECRHS study for their time and cooperation. The EPIC-Norfolk study team are grateful to all the participants who have been part of the EPIC-Norfolk project and to the many members of the study teams at the University of Cambridge who have enabled this research. Generation Scotland is grateful to all the families who took part, the general practitioners and the Scottish School of Primary Care for their help in recruiting them, and the whole Generation Scotland team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, healthcare assistants and nurses. The HUNT study team are grateful for the contributions from He Zhang and Hyun Min Kang and would also like to acknowledge the support given to them by the Genotyping core and Jin Chen. We thank the LBC1936 participants and team members who contributed to this study. The ORCADES study would like to acknowledge the invaluable contributions of the research nurses in Shetland, the administrative team in Edinburgh and the people of Shetland. The VIKING study would like to acknowledge the invaluable contributions of the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney. The Viking Health Study – Shetland (VIKING) DNA extractions and genotyping were performed at the Edinburgh Clinical Research Facility, University of Edinburgh. The Orkney Complex Disease Study (ORCADES) DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. The Raine study would like to acknowledge the continued contribution of Raine Study participants and their families, Raine Study team for cohort coordination and data collection, NHMRC for long term funding over last 30 years, The University of Western Australia, Curtin University, Women and Infants Research Foundation, Telethon Kids Institute, Edith Cowan University, Murdoch University, The University of Notre Dame Australia, and The Raine Medical Research Foundation for providing funding for Core Management of the Raine Study. The Raine study would also like to acknowledge The University of Western Australia (Division of Obstetrics and Gynaecology, King Edward Memorial Hospital and Medical School, Royal Perth Hospital), and Telethon Kids Institute for providing in-kind support for the storage and curation of biological samples, and Pawsey Supercomputing Centre with funding from Australian Government and the Government of Western Australia for providing computation resource to carry out analyses required.
References


Page 14 of 19

Wellcome Open Research 2020, 5:111 Last updated: 25 MAY 2021
Open Peer Review

Current Peer Review Status: ✔ ✔

Version 1

Reviewer Report 05 March 2021

https://doi.org/10.21956/wellcomeopenres.17382.r42679

© 2021 Boateng E. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Eistine Boateng
Early Life Origins of Chronic Lung Diseases, Research Center Borstel, Leibniz Lung Center, Member of the German Center for Lung Research (DZL), Borstel, Germany

In this study, the authors attempted to explain genotype-by-sex interaction on lung function. The membrane protein, hedgehog interacting protein (HHIP), is reported as a susceptibility factor for COPD. Thus, SNPs upstream regulate the expression of HHIP, which is evidently decreased in COPD tissues as shown in related studies. The manuscript is well written and findings could serve as a fundamental basis for future experimental studies. However, I have a few comments which could be considered to improve the impact of the study.

Comments: One key finding in this study was increased expression of (HHIP) in lung tissues from females compared to those from males. Authors should further explain or speculate possible reasons for this observation.

The level of HHIP is known to be decreased in lung tissues from COPD patients. How could the results of this interaction study translate to the molecular pathogenesis of lung diseases eg. COPD vis-à-vis its prevalence in males and females at the study sites where samples were obtained? Again, the development of COPD is characterized by different stages. What is the relevance of the study to staging of lung diseases between males and females?

Lung function partly reflects on the biological state of the organ and authors appear to propose that HHIP may exhibit sex differential effects on lung function. Could authors add a brief outlook for future experimental studies which may want to follow up on their findings? For example, are there differences in cell-specific expression of HHIP in the lungs of males and females and how can this relate to the pathogenesis and risk of lung diseases between the sexes?

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

**Are sufficient details of methods and analysis provided to allow replication by others?**
Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**
I cannot comment. A qualified statistician is required.

**Are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions drawn adequately supported by the results?**
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Pulmonary fibrosis, COPD, asthma, lung development, and miRNA

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

---

Author Response 18 May 2021

**Katherine Fawcett**, University of Leicester, Leicester, UK

We are very grateful to the reviewer for taking the time to read and comment on our manuscript. We will address the comments on a point-by-point basis:

*One key finding in this study was increased expression of (HHIP) in lung tissues from females compared to those from males. Authors should further explain or speculate possible reasons for this observation.*

The higher expression of *HHIP* in females is intriguing, but it is not clear (from our reading of the literature) why this might be.

*The level of HHIP is known to be decreased in lung tissues from COPD patients. How could the results of this interaction study translate to the molecular pathogenesis of lung diseases eg. COPD vis-à-vis its prevalence in males and females at the study sites where samples were obtained? Again, the development of COPD is characterized by different stages. What is the relevance of the study to staging of lung diseases between males and females?*

In our study, the lung-function-lowering allele of rs7697189 was associated with increased expression of *HHIP*. This is the opposite direction of effect to that found in the Zhou et al. paper cited in the discussion, but is consistent with other reports (van der Plaat DA, de Jong K, Lahousse L, Faiz A, Vonk JM, van Diemen CC, Nedeljkovic I, Amin N, Brusselle GG, Hofman A, Brandsma CA, Bossé Y, Sin DD, Nickle DC, van Duijn CM, Postma DS, Boezen HM. Genome-wide association study on the FEV1/FVC ratio in never-smokers identifies HHIP and FAM13A. J Allergy Clin Immunol. 2017 Feb;139(2):533-540. doi: 10.1016/j.jaci.2016.06.062.*
Epub 2016 Sep 6. PMID: 27612410; Morrow, Jarrett D et al. “Functional interactors of three genome-wide association study genes are differentially expressed in severe chronic obstructive pulmonary disease lung tissue.” Scientific reports vol. 7 44232. 13 Mar. 2017, doi:10.1038/srep44232; Lamontagne, Maxime et al. “Refining susceptibility loci of chronic obstructive pulmonary disease with lung eqtls.” PloS one vol. 8,7 e70220. 30 Jul. 2013, doi:10.1371/journal.pone.0070220) that show COPD risk alleles associated with increased HHIP expression. It is possible that, given males have lower expression of HHIP than females, studies of COPD with a higher proportion of males amongst cases than controls could find spurious associations between lower HHIP expression and COPD.

We assume that by "staging of lung diseases" the reviewer is referring to the different trajectories of COPD as outlined, for example, by Lange et al. (Lange P, Celli B, Agustí A, Boje Jensen G, Divo M, Faner R, Guerra S, Marott JL, Martinez FD, Martinez-Cambior P, Meek P, Owen CA, Petersen H, Pinto-Plata V, Schnohr P, Sood A, Soriano JB, Tesfaigzi Y, Vestbo J. Lung-Function Trajectories Leading to Chronic Obstructive Pulmonary Disease. N Engl J Med. 2015 Jul 9;373(2):111-22. doi: 10.1056/NEJMoa1411532. PMID: 26154786). Without understanding the mechanism by which HHIP affects lung function, it is not clear how it would impact trajectories of lung disease. In our discussion, we suggest longitudinal studies to test the effects of HHIP SNPs on trajectories of lung function decline in males and females.

"As UK Biobank participants were aged between 40 and 69 years at recruitment, we did not have the longitudinal data to investigate the effect of HHIP SNPs on trajectories of lung function decline throughout life (28), but this could be an interesting area for future studies."

"Lung function partly reflects on the biological state of the organ and authors appear to propose that HHIP may exhibit sex differential effects on lung function. Could authors add a brief outlook for future experimental studies which might want to follow up on their findings? For example, are there differences in cell-specific expression of HHIP in the lungs of males and females and how can this relate to the pathogenesis and risk of lung diseases between the sexes?"

We do suggest a number of possible avenues for follow up. Though we find no interaction between HHIP SNPs and sex on HHIP expression, our sample size for the expression analysis was quite small. Testing for interaction in larger datasets would therefore be warranted. We also suggest investigating sex differential effects of HHIP at different developmental stages and on trajectories of lung function in longitudinal studies. Finally, we propose testing our in silico-generated hypothesis that "rs7697189 and a SNP in linkage disequilibrium (rs12504628) change transcription factor motifs that may be relevant to the effect of sex hormones on lung development". We are not aware of any cell-type specific datasets that would allow comparison of HHIP expression between males and females, but if these were available then they could indeed throw light on the mechanism of action of sex-differential effects of HHIP. We have added the following to the discussion:

"It is also possible that sex-differential effects of HHIP SNPs are only detectable in particular cell types. We therefore propose that HHIP eQTLs could be tested in larger numbers of males and females and in different cell types."
Competing Interests: No competing interests were disclosed.

Reviewer Report 21 July 2020

https://doi.org/10.21956/wellcomeopenres.17382.r39582

© 2020 Mannino D. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

David M. Mannino

Department of Preventive Medicine and Environmental Health, University of Kentucky College of Public Health, Lexington, KY, USA

The authors provide an interesting analysis of the gene/sex interaction affect on lung function as demonstrated in the UK Biobank cohort and validated in the Spirometa consortium. They found a greater affect in males than females (28 mL vs 9 mL).

Comments: The authors note that this gene is also related to height. Although they adjusted for standing height (and height squared) in the analysis they should note in the limitations that lung function is actually more closely related to sitting height (or even better- thoracic height) which is not well measured. Thus - it is possible that the difference could be explained by other factors related to how we estimate "normal" lung function.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: I am an employee of GlaxoSmithKline
Reviewer Expertise: Pulmonary function, COPD

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 18 May 2021

Katherine Fawcett, University of Leicester, Leicester, UK

We would like to thank the reviewer for taking the time to review our manuscript. We agree that sitting and thoracic height are more related to lung function than standing height and have therefore added the following to the discussion:

"It should be noted, however, that lung function is more closely related to sitting and thoracic height than standing height. We conducted sensitivity analyses showing that the rs7697189-by-sex interaction remained after adjustment for sitting height, but thoracic height was not available."

Competing Interests: No competing interests were disclosed.