RESEARCH ARTICLE

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

Katherine A. Fawcett[1], Ma'en Obeidat[2], Carl Melbourne[1], Nick Shrine[1], Anna L. Guyatt[1], Catherine John[1], Jian'an Luan[3], Anne Richmond[4], Marta R. Moksnes[5], Raquel Granell[6], Stefan Weiss[7], Medea Imboden[8,9], Sebastian May-Wilson[10], Pirro Hysi[11], Thibaud S. Boutin[4], Laura Portas[12], Claudia Flexeder[13], Sarah E. Harris[14,15], Carol A. Wang[16], Leo-Pekka Lyytikäinen[17-19], Teemu Palviainen[20], Rachel E. Foong[21,22], Dirk Keidel[8,9], Cosetta Minelli[12], Claudia Langenberg[3], Yohan Bossé[23], Maarten Van den Berge[24], Don D. Sin[2,25], Ke Hao[26], Archie Campbell[27], David Porteous[27], Sandosh Padmanabhan[28], Blair H. Smith[29], David M. Evans[6,30,31], Sue Ring[6,30], Arnulf Langhammer[32], Kristian Hveem[5], Cristen Willer[33-35], Ralf Ewert[36], Beate Stubbe[36], Nicola Pirastu[10], Lucija Klaric[4], Peter K. Joshi[10], Karina Patasova[11], Mangino Massimo[11], Ozren Polasek[37], John M. Starr[14,38+], Stefan Karrasch[39-41], Konstantin Strauch[42,43], Thomas Meitinger[44,45], Igor Rudan[10], Taina Rantanen[46], Kirsi Pietiläinen[47,48], Mika Kähönen[49,50], Olli T. Raitakari[51-53], Graham L. Hall[21,22], Peter D. Sly[54], Craig E. Pennell[16], Jaakko Kaprio[20,55], Terho Lehtimäki[17,18], Veronique Vitart[4], Ian J. Deary[14,15], Debbie Jarvis[12,56], James F. Wilson[4,10], Tim Spector[11], Nicole Probst-Hensch[8,9], Nicholas J. Wareham[3], Henry Völzke[57], John Henderson[30+], David P. Strachan[58], Ben M. Brumpton[5,59,60], Caroline Hayward[4], Ian P. Hall[61], Martin D. Tobin[1,62], Louise V. Wain[1,62]

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

[11]The Department of Twin Research & Genetic Epidemiology, King's College London, St Thomas' Campus, Lambeth Palace Road, London, UK
[12]Population Health and Occupational Disease, National Heart and Lung Institute, Imperial College London, London, UK
[13]Institute of Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, 85764, Germany
[14]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, EH8 9JZ, UK
[15]Psychology, University of Edinburgh, Edinburgh, EH8 9JZ, UK
[16]School of Medicine and Public Health, Faculty of Medicine and Health, The University of Newcastle, Callaghan, Australia
[17]Department of Clinical Chemistry, Fimlab Laboratories, Tampere, 33520, Finland
[18]Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, 33014, Finland
[19]Department of Cardiology, Heart Center, Tampere University Hospital, Tampere, 33521, Finland
[20]Institute for Molecular Medicine FIMM, University of Helsinki, Helsinki, FI-00014, Finland
[21]Telethon Kids Institute, Perth, Australia
[22]School of Physiotherapy and Exercise Science, Faculty of Health Sciences, Curtin University, Perth, Australia
[23]Institut universitaire de cardiologie et de pneumologie de Québec, Department of Molecular Medicine, Laval University, Québec, Canada
[24]University Medical Center Groningen, Department of Pulmonology, GRIAC Research Institute, University of Groningen, Groningen, The Netherlands
[25]Respiratory Division, Department of Medicine, University of British Columbia, Vancouver, BC, Canada
[26]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, USA
[27]Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, EH4 2XU, UK
[28]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
[29]Division of Population Health Sciences, Ninewells Hospital and Medical School, University of Dundee, Dundee, DD1 9SY, UK
[30]Population Health Sciences Bristol Medical School, University of Bristol, Bristol, BS8 2BN, UK
[31]University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, QLD 4072, Australia
[32]Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology, Trondheim, Norway
[33]Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, USA
[34]Department of Internal Medicine, University of Michigan, Ann Arbor, USA
[35]Department of Human Genetics, University of Michigan, Ann Arbor, USA
[36]Department of Internal Medicine B, Cardiology, Pneumology, Infectious Diseases, Intensive Care Medicine, University Medicine Greifswald, Greifswald, 17475, Germany
[37]University of Split School of Medicine, Split, Croatia
[38]Alzheimer Scotland Research Centre, University of Edinburgh, Edinburgh, EH8 9JZ, UK
[39]Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, Ludwig-Maximilians-Universität, Munich, 80336, Germany
[40]Institute of Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, 85764, Germany
[41]Comprehensive Pneumology Center Munich (CPC-M), Member of the German Center for Lung Research (DZL), Munich, 81377, Germany
[42]Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany
[43]Chair of Genetic Epidemiology, IBE, Faculty of Medicine, LMU Munich, Munich, 81377, Germany
[44]Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany
[45]Institute of Human Genetics, Klinikum rechts der Isar der TU Muenchen, Muenchen, 81675, Germany
[46]Faculty of Sport and Health Sciences, Gerontology Research Center, University of Jyväskylä, Jyväskylä, Finland
[47]Obesity Research Unit, Research Program for Clinical and Molecular Metabolism, Faculty of Medicine, University of Helsinki, Helsinki, FI-00014, Finland
[48]Obesity Centre, Abdominal Centre, Helsinki University Hospital and University of Helsinki, Helsinki, FI-00029, Finland
[49]Department of Clinical Physiology, Tampere University Hospital, Tampere, 33521, Finland
[50]Department of Clinical Physiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, 33014, Finland
[51]

Centre for Population Health Research, University of Turku and Turku University Hospital, Turku, Finland
[52]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland
[53]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland
[54]Children's Health and Environment Program, The University of Queensland, Brisbane, Australia
[55]Department of Public Health, University of Helsinki, Helsinki, FI-00014, Finland
[56]MRC-PHE Centre for the Environment and Health, London, UK
[57]Intitute for Community Medicine, University Medicine Greifswald, Greifswald, 17487, Germany
[58]Population Health Research Institute, St George's, University of London, London, SW17 0RE, UK
[59]Clinic of Thoracic and Occupational Medicine, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway
[60]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
[61]Division of Respiratory Medicine and NIHR-Nottingham Biomedical Research Centre, University of Nottingham, Nottingham, UK
[62]National Institute for Health Research, Leicester Respiratory Biomedical Research Centre, Glenfield Hospital, Leicester, LE3 9QP, UK

[+] Deceased author

v1 **First published:** 01 Jun 2020, **5**:111
https://doi.org/10.12688/wellcomeopenres.15846.1
Latest published: 01 Jun 2020, **5**:111
https://doi.org/10.12688/wellcomeopenres.15846.1

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_1) ($P=3.15x10^{-15}$), was replicated (P=0.016) in SpiroMeta. The C allele increased FEV_1 more in males (untransformed FEV_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

Open Peer Review

Reviewer Status ✓ ✓

	Invited Reviewers	
	1	**2**
version 1 01 Jun 2020	✓ report	✓ report

1. **David M. Mannino**, University of Kentucky College of Public Health, Lexington, USA

2. **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

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding authors: Katherine A. Fawcett (kaf19@leicester.ac.uk), Louise V. Wain (lvw1@leicester.ac.uk)

Author roles: Fawcett KA: Conceptualization, Formal Analysis, Investigation, Methodology, Project Administration, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Obeidat M**: Formal Analysis, Investigation, Resources, Visualization, Writing – Review & Editing; **Melbourne C**: Conceptualization, Methodology, Software, Writing – Original Draft Preparation, Writing – Review & Editing; **Shrine N**: Conceptualization, Data Curation, Methodology, Resources, Software, Writing – Original Draft Preparation, Writing – Review & Editing; **Guyatt AL**: Formal Analysis, Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **John C**: Formal Analysis, Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **Luan J**: Formal Analysis, Investigation, Writing – Review & Editing; **Richmond A**: Formal Analysis, Investigation, Writing – Review & Editing; **Moksnes MR** : Formal Analysis, Investigation, Writing – Review & Editing; **Granell R**: Formal Analysis, Investigation, Resources, Writing – Review & Editing; **Weiss S**: Formal Analysis, Investigation, Resources, Writing – Review & Editing; **Imboden M**: Conceptualization, Formal Analysis, Investigation, Resources, Writing – Review & Editing; **May-Wilson S**: Formal Analysis, Investigation, Writing – Review & Editing; **Hysi P**: Conceptualization, Resources, Writing – Review & Editing; **Boutin TS**: Formal Analysis, Investigation, Writing – Review & Editing; **Portas L**: Formal Analysis, Investigation, Writing – Review & Editing; **Flexeder C**: Formal Analysis, Investigation, Writing – Review & Editing; **Harris SE**: Formal Analysis, Investigation, Resources, Writing – Review & Editing; **Wang CA**: Formal Analysis, Investigation, Resources, Writing – Review & Editing; **Lyytikäinen LP**: Formal Analysis, Investigation, Resources, Writing – Review & Editing; **Palviainen T**: Formal Analysis, Investigation, Writing – Review & Editing; **Foong RE**: Formal Analysis, Investigation, Resources, Writing – Review & Editing; **Keidel D**: Formal Analysis, Investigation, Resources, Writing – Review & Editing; **Minelli C**: Resources, Writing – Review & Editing; **Langenberg C**: Conceptualization, Resources, Writing – Review & Editing; **Bossé Y**: Resources, Writing – Review & Editing; **Van den Berge M**: Formal Analysis, Investigation, Writing – Review & Editing; **Sin DD**: Conceptualization, Writing – Review & Editing; **Hao K**: Conceptualization, Writing – Review & Editing; **Campbell A**: Resources, Writing – Review & Editing; **Porteous D**: Resources, Writing – Review & Editing; **Padmanabhan S**: Resources, Writing – Review & Editing; **Smith BH**: Resources, Writing – Review & Editing; **Evans DM**: Resources, Writing – Review & Editing; **Ring S**: Resources, Writing – Review & Editing; **Langhammer A**: Resources, Writing – Review & Editing; **Hveem K**: Resources, Writing – Review & Editing; **Willer C**: Resources, Writing – Review & Editing; **Ewert R**: Conceptualization, Resources, Writing – Review & Editing; **Stubbe B**: Conceptualization, Resources, Writing – Review & Editing; **Pirastu N**: Formal Analysis, Investigation, Writing – Review & Editing; **Klaric L**: Resources, Writing – Review & Editing; **Joshi PK**: Resources, Writing – Review & Editing; **Patasova K**: Formal Analysis, Investigation, Writing – Review & Editing; **Massimo M**: Resources, Writing – Review & Editing; **Polasek O**: Conceptualization, Resources, Writing – Review & Editing; **Starr JM**: Conceptualization, Resources, Writing – Review & Editing; **Karrasch S**: Conceptualization, Resources, Writing – Review & Editing; **Strauch K**: Resources, Writing – Review & Editing; **Meitinger T**: Resources, Writing – Review & Editing; **Rudan I**: Conceptualization, Resources, Writing – Review & Editing; **Rantanen T**: Resources, Writing – Review & Editing; **Pietiläinen K**: Conceptualization, Resources, Writing – Review & Editing; **Kähönen M**: Conceptualization, Resources, Writing – Review & Editing; **Raitakari OT**: Conceptualization, Resources, Writing – Review & Editing; **Hall GL**: Conceptualization, Resources, Writing – Review & Editing; **Sly PD**: Conceptualization, Resources, Writing – Review & Editing; **Pennell CE**: Conceptualization, Resources, Writing – Review & Editing; **Kaprio J**: Conceptualization, Resources, Writing – Review & Editing; **Lehtimäki T**: Conceptualization, Resources, Writing – Review & Editing; **Vitart V**: Resources, Writing – Review & Editing; **Deary IJ**: Conceptualization, Resources, Writing – Review & Editing; **Jarvis D**: Resources, Writing – Review & Editing; **Wilson JF**: Conceptualization, Resources, Writing – Review & Editing; **Spector T**: Resources, Writing – Review & Editing; **Probst-Hensch N**: Conceptualization, Formal Analysis, Investigation, Resources, Writing – Review & Editing; **Wareham NJ**: Conceptualization, Formal Analysis, Investigation, Resources, Writing – Review & Editing; **Völzke H**: Conceptualization, Resources, Writing – Review & Editing; **Henderson J**: Resources, Writing – Review & Editing; **Strachan DP**: Conceptualization, Formal Analysis, Investigation, Resources, Writing – Review & Editing; **Brumpton BM**: Conceptualization, Formal Analysis, Investigation, Resources, Writing – Review & Editing; **Hayward C**: Conceptualization, Resources, Writing – Review & Editing; **Hall IP**: Conceptualization, Data Curation, Funding Acquisition, Methodology, Project Administration, Resources, Supervision, Writing – Review & Editing; **Tobin MD**: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Wain LV**: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

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

the work of the authors and will serve as guarantors for the contents of this paper. CROATIA-Korcula/Split/Vis: MRC, University Unit Programme Grant (MC_PC_U127592696), European Union, EUROSPAN project (contract no. LSHG-CT-2006-018947), Croatian Ministry of Science (216-1080315-0302) (I.R.) ECRHS: This work was supported by a contract from the European Commission (018996), Fondo de Investigación Sanitaria (91/0016-060-05/E, 92/0319, 93/0393, 97/0035-01, 99/0034-01 and 99/0034-02), Hospital General de Albacete, Hospital General Ramón Jiménez, Consejería de Sanidad del Principado de Asturias, CIRIT (1997SGR 00079, 1999SGR 00241), and Servicio Andaluz de Salud, SEPAR, Public Health Service (R01 HL62633-01), RCESP (C03/09), Red RESPIRA (C03/011), Basque Health Department, Swiss National Science Foundation, Swiss Federal Office for Education and Science, Swiss National Accident Insurance Fund (SUVA), GSF-National Research Centre for Environment and Health, Deutsche Forschungsgemeinschaft (DFG) (FR 1526/1-1, MA 711/4-1), Programme Hospitalier de Recherche Clinique-DRC de Grenoble 2000 no. 2610, Ministry of Health, Direction de la Recherche Clinique, Ministere de l'Emploi et de la Solidarite, Direction Generale de la Sante, CHU de Grenoble, Comite des Maladies Respiratoires de l'Isere. UCB-Pharma (France), Aventis (France), Glaxo France. Estonian Science Foundation, and Asthma UK (formerly known as National Asthma Campaign UK). EPIC-Norfolk: Medical Research Council (MR/N003284/1, MC-UU_12015/1, MC_PC_13048) Cancer Research UK (C864/A14136). FTC: Academy of Finland (308248, 312073) (J.K.), Sigrid Juselius Foundation (J.K.), Academy of Finland (213506) (T.R.), Academy of Finland (272376, 314383, 266286) (K.P.), Finnish Medical Foundation (K.P.), Novo Nordisk Foundation (K.P.), Finnish Diabetes Research Foundation (K.P.), State Research Funds (K.P.), University of Helsinki (K.P.) Generation Scotland: MRC, University Unit Programme Grant (MC_PC_U127592696), Wellcome Trust Strategic Award (104036), Chief Scientist Office (CZD/16/6), Scottish Funding Council (HR03006). HUNT: Stiftelsen Kristian Gerhard Jebsen (K.H.), The Liaison Committee for education, research and innovation in Central Norway (K.H., B.M.B.), NIH (HL135824, HL109946, HL127564) (C.W.): The Nord-Trøndelag Health Study (The HUNT Study) is a collaboration between HUNT Research Center (Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology), Nord-Trøndelag County Council, Central Norway Regional Health Authority, and the Norwegian Institute of Public Health. B.M.B. received a research grant from the Liaison Committee between the Central Norway Regional Health Authority and the Norwegian University of Science and Technology. KORA: MC-Health, LMUinnovativ: The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. LBC1936: Biotechnology and Biological Sciences Research Council (BBSRC) (BB/F019394/1), Age UK (The Disconnected Mind Project) (DCM and DCM PHASE 2), Cross Council Lifelong Health and Wellbeing Initiative (MR/K026992/1). ORCADES/VIKING: Chief Scientist Office (CZB/4/276, CZB/4/710) (J.F.W.), MRC (MC_UU_00007/10) (J.F.W.), MRC (MR/N013166/1) (S.M.), EU FP6 (LSHG-CT-2006-018947) (J.F.W.), Royal Society (URF to J.F.W.). The work of LK was supported by an RCUK Innovation Fellowship from the National Productivity Investment Fund (MR/R026408/1). The Raine Study: National Health and Medical Research Council of Australia (NHMRC) (572613, 403981, 003209), Canadian Institutes of Health Research (CIHR) (MOP-82893), Raine Medical Research Foundation. SAPALDIA: Swiss National Science Foundation (33CS30-148470/1&2, 33CSCO-134276/1, 33CSCO-108796, 324730_135673, 3247BO-104283, 3247BO-104288, 3247BO-104284, 3247-065896, 3100-059302, 3200-052720, 3200-042532, 4026-028099, PMPDP3_129021/1, PMPDP3_141671/1) (SAPALDIA1 to SAPALIDA5), Canton's government of Aargau, Basel-Stadt, Basel-Land, Geneva, Luzern, Ticino, Valais, and Zürich, Federal Offices of Environment, of Public Health, and of Roads and Transport, Cantonal lung leages of Basel Stadt/ Basel Landschaft, Geneva, Ticino, Valais, Graubünden and Zurich, Horizon 2020 (633212) (ALEC Project), European Commission (308610, 018996) (GABRIEL Project), Freie Akademische Gesellschaft (N.P.), UBS Wealth Foundation (N.P.), Wellcome Trust (084703) (GABRIEL Project), Talecris Biotherapeutics GmbH (N.P.), Abbott Diagnostics (N.P.) SHIP/SHIP_Trend/SHIP_Trend_B2: German Federal Ministry of Education and Research (01ZZ9603, 01ZZ0103, and 01ZZ0403), German Research Foundation (GR 1912/5-1) YFS: Academy of Finland (286284, 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi)), Social Insurance Institution of Finland, Competitive State Research Financing of the Expert Responsibility area of Kuopio, Tampere and Turku University Hospitals, Juho Vainio Foundation, Paavo Nurmi Foundation, Finnish Foundation for Cardiovascular Research, Finnish Cultural Foundation, The Sigrid Juselius Foundation, Tampere Tuberculosis Foundation, Emil Aaltonen Foundation, Yrjö Jahnsson Foundation, Signe and Ane Gyllenberg Foundation, Diabetes Research Foundation of Finnish Diabetes Association, EU Horizon 2020 (755320 for TAXINOMISIS), European Research Council (742927 for MULTIEPIGEN), Tampere University Hospital Supporting Foundation.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Fawcett KA, Obeidat M, Melbourne C *et al.* **Variants associated with *HHIP* expression have sex-differential effects on lung function [version 1; peer review: 2 approved]** Wellcome Open Research 2020, **5**:111 https://doi.org/10.12688/wellcomeopenres.15846.1

First published: 01 Jun 2020, **5**:111 https://doi.org/10.12688/wellcomeopenres.15846.1

Introduction

Measures of lung function, including forced expiratory volume in 1 second (FEV_1) 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[1]. 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[2]. It is known that sex hormones can influence these processes but the mechanisms are not well understood[3,4]. 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[5,6]. 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[5].

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[7]. 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[8]. 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[9].

In this study, we tested for an interaction effect of 7,745,864 variants and sex on FEV_1, FEV_1/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) $\geq$ 0.01 and imputation quality scores $\geq$ 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 ($P<5 \times 10^{-8}$) interaction with sex on at least one of four lung function traits (FEV_1, FEV_1/FVC, FVC, and PEF) with a further 21 SNPs showing suggestive significance ($P<1 \times 10^{-6}$) (Table 1; Figure S1, *Extended data*[10]). 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 $P = 3.15 \times 10^{-15}$), rs9403386 near the gene encoding Adhesion G Protein-Coupled Receptor G6 (*ADGRG6*, previously known as *GPR126*) (interaction $P = 4.56 \times 10^{-9}$), and rs162185 downstream of the gene encoding transcription factor 21 (*TCF21*) (interaction $P = 4.87 \times 10^{-9}$)[11–16]. This may, in part, reflect greater power to detect interactions with variants with strong main effects on lung function. Only rs355079 (interaction $P = 8.84 \times 10^{-7}$) 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 r^2 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*[10]). 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 ($\beta = 0.052$ [SE 0.004], $P = 2.13 \times 10^{-33}$) compared to females ($\beta = 0.013$ [SE 0.003], $P = 1.16 \times 10^{-5}$) (Table 1). This SNP resides upstream of the *HHIP* gene and is in linkage disequilibrium with two previously reported lung function-associated sentinel SNPs, rs13141641[16,17] ($r^2 = 0.91$) and rs13116999[17] ($r^2 = 0.56$). SNP rs7697189 also showed some evidence of interaction with sex on PEF ($\beta = -0.035$ (0.005), $P = 8.78 \times 10^{-12}$), FEV_1/FVC ($\beta = -0.028$ (0.005), $P = 8.98 \times 10^{-8}$), and FVC ($\beta = -0.020$ (0.005), $P = 8.71 \times 10^{-5}$) (Table S1, *Extended data*[10]; Figure 2).

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[18] 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 FEV_1 was consistent with the original estimate after adjustment for sitting height ($\beta = -0.04$ [SE = 0.005], $P = 1.97 \times 10^{-15}$).

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.

SNP (nearest gene) and coordinates	Test/other allele	Trait	Lung function UK Biobank males			Lung function UK Biobank females			Sex interaction in UK Biobank		Sex interaction in SpiroMeta	
			MAF	Beta (SE)	P	MAF	Beta (SE)	P	Beta (SE)	P	Beta (SE)	P
rs7697189 (HHIP) 4:145436894	C/G	FEV_1	0.390	0.052 (0.004)	2.13E-33	0.392	0.013 (0.003)	1.16E-05	-0.040 (0.005)	3.15E-15	**-0.025 (0.01)**	**0.016**
rs9403386 (ADGRG6) 6:142764073	C/A	FEV_1/FVC	0.031	0.214 (0.012)	4.48E-75	0.031	0.128 (0.009)	2.16E-43	-0.086 (0.015)	4.56E-09	**-0.035 (0.032)**	**0.281**
rs162185 (TCF21) 6:134226147	C/T	PEF	0.411	-0.038 (0.004)	1.35E-18	0.410	-0.009 (0.003)	0.002	0.030 (0.005)	4.87E-09	**0.022 (0.0139)**	**0.083**
rs6480592 (CHST3) 10:73764509	C/T	PEF	0.398	-0.021 (0.004)	1.66E-06	0.400	0.007 (0.003)	0.011	0.028 (0.005)	2.85E-08	**0.003 (0.012)**	**0.808**
rs111893604 (ZSCAN10) 16:3141104	G/T	FEV_1	0.059	0.040 (0.009)	1.70E-05	0.059	-0.020 (0.006)	0.002	-0.060 (0.011)	4.04E-08	0.006 (0.026)	0.827
rs72694266 (RP11-907D1.1) 14:97578576	A/C	PEF	0.077	-0.044 (0.008)	2.69E-07	0.078	0.008 (0.006)	0.145	0.053 (0.010)	6.31E-08	-0.049 (0.027)	0.066
rs72781459 10:10247676	C/T	PEF	0.096	0.031 (0.007)	3.44E-05	0.097	-0.012 (0.005)	0.014	-0.046 (0.009)	1.08E-07	0.007 (0.021)	0.729
rs74316059 (RP11-649A16.1) 3:146983325	T/C	FEV_1/FVC	0.042	0.049 (0.010)	2.52E-06	0.043	-0.018 (0.008)	0.029	-0.068 (0.013)	2.38E-07	**-0.031 (0.028)**	**0.269**
rs55789572 (EIF2S2/RALY) 20:32687822	A/C	FEV_1	0.022	0.041 (0.015)	0.006	0.022	-0.047 (0.010)	2.67E-06	-0.089 (0.017)	2.80E-07	**-0.01 (0.033)**	**0.765**
rs74933518 (DAPK2) 15:64303295	A/G	PEF	0.025	-0.072 (0.014)	1.23E-07	0.025	0.007 (0.009)	0.421	0.082 (0.016)	3.05E-07	**0.025 (0.043)**	**0.568**
rs11247571 (ABR) 17:908502	G/A	PEF	0.343	-0.025 (0.005)	3.65E-08	0.344	0.002 (0.003)	0.569	0.027 (0.005)	3.22E-07	**0.010 (0.014)**	**0.473**
rs707588 (RP11-154H17.1) 1:5711430	G/A	FEV_1	0.482	-0.020 (0.004)	3.23E-06	0.482	0.006 (0.003)	0.029	0.025 (0.005)	3.27E-07	**0.014 (0.01)**	**0.183**
rs138473298 (AUTS2) 7:69644989	T/C	PEF	0.012	-0.077 (0.020)	0.0002	0.011	0.043 (0.014)	0.002	0.122 (0.024)	3.52E-07	**0.037 (0.060)**	**0.540**
rs139069254 (RP11-648K4.2) 15:88113916	A/G	FEV_1	0.018	0.071 (0.016)	1.83E-05	0.018	-0.027 (0.011)	0.017	-0.098 (0.019)	4.66E-07	**-0.051 (0.041)**	**0.216**

Wellcome Open Research 2020, 5:111 Last updated: 05 MAR 2021

SNP (nearest gene) and coordinates	Test/other allele	Trait	Lung function UK Biobank males			Lung function UK Biobank females			Sex interaction in UK Biobank		Sex interaction in SpiroMeta	
			MAF	Beta (SE)	P	MAF	Beta (SE)	P	Beta (SE)	P	Beta (SE)	P
rs138163836 (PVRL3) 3:110952902	C/T	FVC	0.021	0.064 (0.015)	1.94E-05	0.020	-0.025 (0.011)	0.019	-0.091 (0.018)	5.07E-07	**-0.025 (0.038)**	**0.5**
rs28493055 (XDH) 2:31573390	T/G	FEV_1	0.012	0.065 (0.020)	0.002	0.013	-0.055 (0.014)	6.40E-05	-0.119 (0.024)	5.60E-07	0.035 (0.054)	0.519
rs117380804 18:76145905	T/C	FVC	0.035	0.035 (0.012)	0.003	0.036	-0.035 (0.008)	1.93E-05	-0.070 (0.014)	6.25E-07	**-0.034 (0.03)**	**0.255**
rs602622 (RASGRP3) 2:33658226	C/G	PEF	0.444	-0.022 (0.004)	2.11E-07	0.445	0.002 (0.003)	0.444	0.025 (0.005)	6.45E-07	-0.013 (0.013)	0.323
rs2253718 (RF00019, SFTA2) 6:30900427	T/G	PEF	0.409	-0.049 (0.004)	5.69E-30	0.405	-0.027 (0.003)	1.78E-20	0.025 (0.005)	7.05E-07	**0.002 (0.016)**	**0.925**
rs2353939 (HHIP) 4:145729724	G/A	FVC	0.437	0.016 (0.004)	0.0002	0.435	-0.009 (0.003)	0.002	-0.025 (0.005)	7.55E-07	**-0.016 (0.01)**	**0.124**
rs7691139 (ZNF280A) 22:22876151	G/C	FEV_1/FVC	0.116	-0.025 (0.007)	0.0003	0.115	0.017 (0.005)	0.002	0.043 (0.009)	7.62E-07	Not tested	
rs13020954 2:17296984	C/T	FEV_1/FVC	0.014	0.050 (0.017)	0.004	0.014	-0.057 (0.014)	3.83E-05	-0.109 (0.022)	7.88E-07	**-0.062 (0.043)**	**0.148**
rs2731120 (MLF1) 3:158297633	A/C	FVC	0.346	0.029 (0.004)	3.72E-11	0.346	0.003 (0.003)	0.310	-0.026 (0.005)	8.14E-07	**-0.008 (0.011)**	**0.433**
rs355079 (LMCD1-AS1) 3:8643371	T/C	FVC	0.337	0.015 (0.004)	0.0007	0.339	-0.011 (0.003)	0.0004	-0.026 (0.005)	8.84E-07	0.001 (0.011)	0.935
rs7338055 (SPRYD7) 13:50504226	C/A	FVC	0.259	0.018 (0.005)	0.0001	0.259	-0.009 (0.003)	0.008	-0.028 (0.006)	9.81E-07	**-0.008 (0.012)**	**0.478**
rs34490170 (NEUROD1/ CERKL) 2:182576419	C/T	FVC	0.110	-0.035 (0.007)	6.41E-07	0.110	0.007 (0.005)	0.186	0.041 (0.008)	9.95E-07	**0.009 (0.018)**	**0.622**

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, age^2, 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 (rs7697189) 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 final column indicates that the effect in SpiroMeta was in the same direction to the effect in UK Biobank.

Wellcome Open Research 2020, 5:111 Last updated: 05 MAR 2021

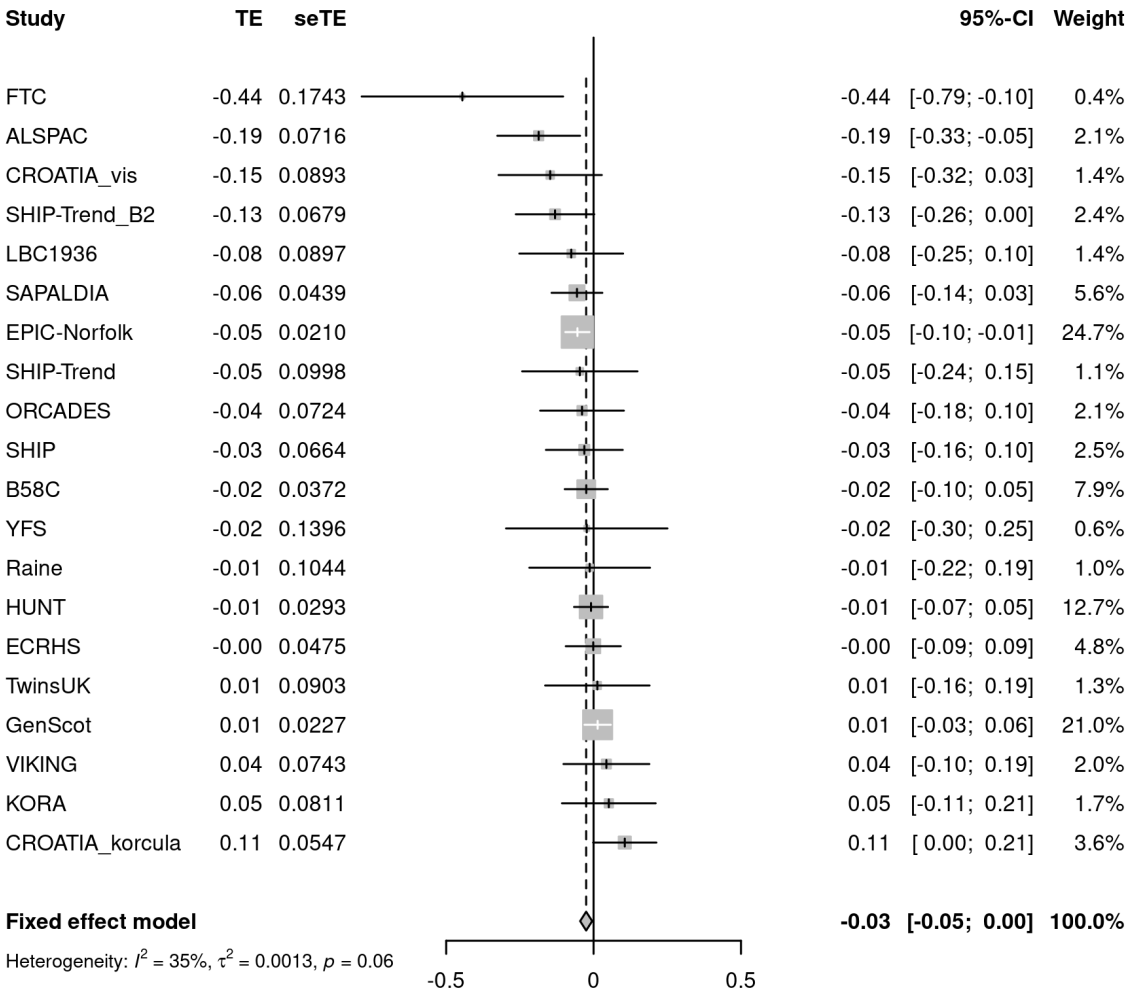

Figure 1. Meta-analysis of rs7697189-by-sex interaction effects on lung function in SpiroMeta cohorts. The forest plot shows the beta-coefficients (test effects, TE) and standard errors for the interaction between rs7697189 and sex on forced expiratory volume in 1 second (FEV_1) in 20 cohorts of the SpiroMeta consortium (total N = 75,696). The overall effect size from fixed effects meta-analysis is represented by the diamond.

lung function in males compared to females could arise if there was an interaction effect with smoking. However, there was no interaction between rs7697189 and ever-smoking status on FEV_1 in this study (interaction P = 0.63). Pack years data was available for 94,750 UK Biobank participants. In sensitivity analyses we found a similar rs7697189-by-sex effect size on FEV_1 when adjusted for pack years (β = -0.033 [SE = 0.009], P = 3.50 x 10^{-4}) and no interaction between genotype and pack years on FEV_1 (interaction P = 0.80).

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

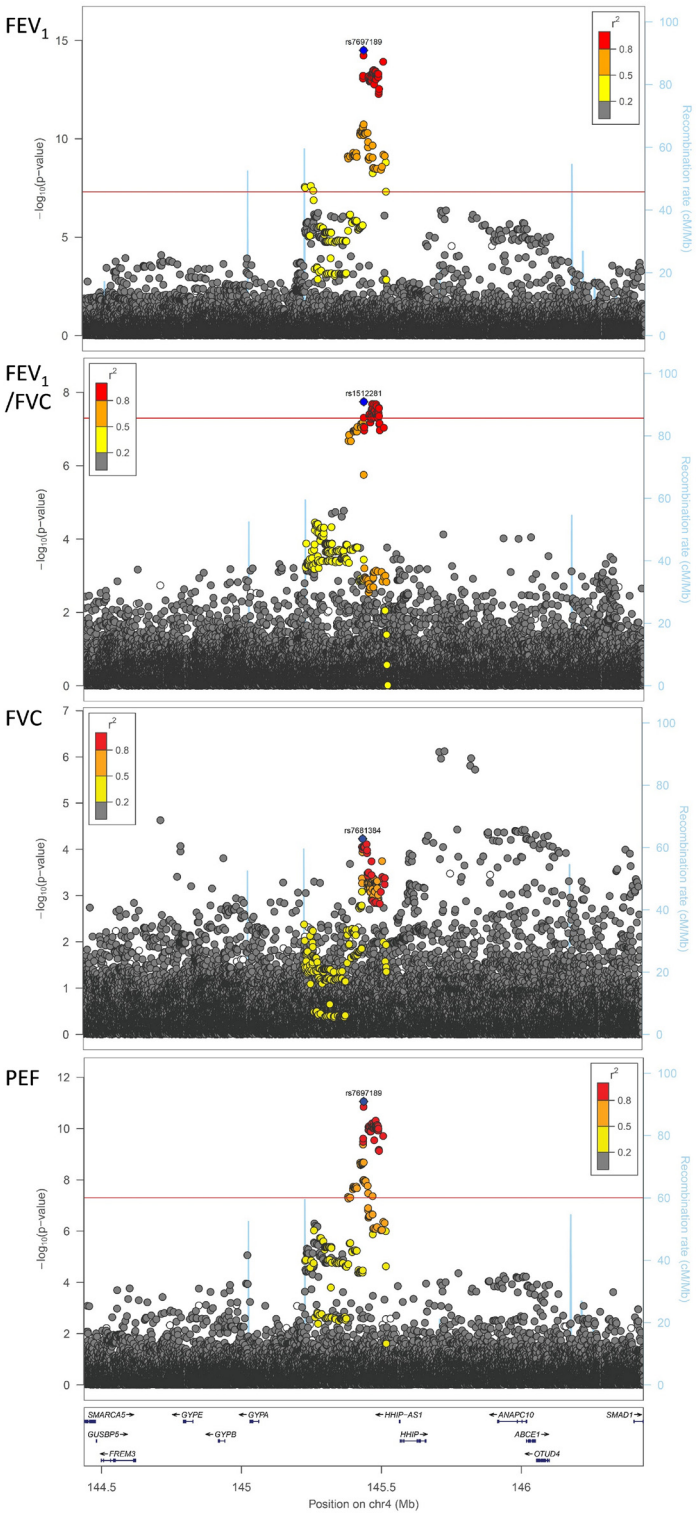

Figure 2. Genotype-by-sex interaction results within the *HHIP* region for lung function traits in UK Biobank. The SNP with the strongest association in the rs7697189-proximal region is represented by a blue diamond. The FEV_1 and PEF sentinels are rs7697189, the FEV_1/FVC sentinel is rs1512281 (R^2 = 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*[10]). 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*[10]). This agrees with GTEx data on HHIP lung expression in males and females (Figure S4, *Extended data*[10]).

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 $P = 3.15 \times 10^{-15}$). The SNP showed some evidence of replication in 75,696 individuals from 20 independent studies of the SpiroMeta consortium (β = -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 β = 0.052 (0.004) in males and 0.013 (0.003) in females, corresponding to an untransformed FEV_1 β = 0.028 [SE 0.0022] litres in males vs β = 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 (β = -0.028 (0.005), $P = 8.78 \times 10^{-12}$ and β = -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[25].

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[19]. 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[19], 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 *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[8].

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*,[10]).

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[7], (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*[10]).

Participants' DNA was genotyped using either the Affymetrix Axiom® UK BiLEVE array or the Affymetrix Axiom® UK Biobank array[29]. Genotypes were imputed based on the Human Reference Consortium (HRC) panel, as described elsewhere[29]. 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*[10]). Ten studies (N=17,280) were imputed using 1000 Genomes Phase 1 reference panel[30,31], nine (N=37,919) were imputed using the Haplotype Reference Consortium (HRC) panel[29], 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*[10]). Measurements of spirometry for each study are as previously described[7,21]. Fourteen SpiroMeta studies had data on PEF (N=51,555).

Statistical analysis

Spirometry-based lung function traits FEV_1, FEV_1/FVC, FVC, and PEF were pre-adjusted for age, age^2, standing height (or sitting height in the sensitivity analysis) and smoking status and the residuals rank-transformed to normality using the rntransform 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[32,33].

Regression analysis to test genotype-by-sex interactions on height were conducted using a model including genotype (additive effect), age, age^2, 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[34]. 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[35–37] and in S1 Appendix (*Extended data*[10]). *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/ 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[38]

Extended data

Figshare: Variants associated with *HHIP* expression have sex-differential effects on lung function: supplementary material. https://doi.org/10.6084/m9.figshare.12129207[10]

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 FEV_1 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

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 SPECTRE 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

1. GBD 2016 Causes of Death Collaborators: **Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016: a systematic analysis for the Global Burden of Disease Study 2016.** *Lancet.* 2017; **390**(10100): 1151–210.
PubMed Abstract | Publisher Full Text | Free Full Text

2. LoMauro A, Aliverti A: **Sex differences in respiratory function.** *Breathe (Sheff).* 2018; **14**(2): 131–40.
PubMed Abstract | Publisher Full Text | Free Full Text

3. Kocurek EG, Hemnes AR: **Women's Health and Lung Development and Disease.** *Obstet Gynecol Clin North Am.* 2016; **43**(2): 307–23.
PubMed Abstract | Publisher Full Text

4. Townsend EA, Miller VM, Prakash YS: **Sex differences and sex steroids in lung health and disease.** *Endocr Rev.* 2012; **33**(1): 1–47.
PubMed Abstract | Publisher Full Text | Free Full Text

5. Aryal S, Diaz-Guzman E, Mannino DM: **COPD and gender differences: an update.** *Transl Res.* 2013; **162**(4): 208–18.
PubMed Abstract | Publisher Full Text

6. Sorheim IC, Johannessen A, Gulsvik A, *et al.*: **Gender differences in COPD: are women more susceptible to smoking effects than men?** *Thorax.* 2010; **65**(6): 480–5.
PubMed Abstract | Publisher Full Text

7. 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; **51**(3): 481–93.
PubMed Abstract | Publisher Full Text | Free Full Text

8. Hardin M, Cho MH, Sharma S, *et al.*: **Sex-Based Genetic Association Study Identifies CELSR1 as a Possible Chronic Obstructive Pulmonary Disease Risk Locus among Women.** *Am J Respir Cell Mol Biol.* 2017; **56**(3): 332–41.
PubMed Abstract | Publisher Full Text | Free Full Text

9. Khramtsova EA, Davis LK, Stranger BE: **The role of sex in the genomics of human complex traits.** *Nat Rev Genet.* 2019; **20**(3): 173–190.
PubMed Abstract | Publisher Full Text

10. Fawcett K, Obeidat M, Melbourne C, *et al.*: **Fawcett_et_al_Extended_data_supplement.docx.** *figshare.* Journal contribution. 2020.
http://www.doi.org/10.6084/m9.figshare.12129207.v1

11. Kichaev G, Bhatia G, Loh PR, *et al.*: **Leveraging Polygenic Functional Enrichment to Improve GWAS Power.** *Am J Hum Genet.* 2019; **104**(1): 65–75.
PubMed Abstract | Publisher Full Text | Free Full Text

12. Pillai SG, Ge D, Zhu G, *et al.*: **A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci.** *PLoS Genet.* 2009; **5**(3): e1000421.
PubMed Abstract | Publisher Full Text | Free Full Text

13. Soler Artigas M, Wain LV, Miller S, *et al.*: **Sixteen new lung function signals identified through 1000 Genomes Project reference panel imputation.** *Nat Commun.* 2015; **6**: 8658.
PubMed Abstract | Publisher Full Text | Free Full Text

14. Terzikhan N, Sun F, Verhamme FM, *et al.*: **Heritability and genome-wide association study of diffusing capacity of the lung.** *Eur Respir J.* 2018; **52**(3): 1800647.
PubMed Abstract | Publisher Full Text

15. Van Durme YM, Eijgelsheim M, Joos GF, *et al.*: **Hedgehog-interacting protein is a COPD susceptibility gene: the Rotterdam Study.** *Eur Respir J.* 2010; **36**(1): 89–95.
PubMed Abstract | Publisher Full Text

16. Wilk JB, Chen TH, Gottlieb DJ, *et al.*: **A genome-wide association study of pulmonary function measures in the Framingham Heart Study.** *PLoS Genet.* 2009; **5**(3): e1000429. two studies of COPD genetics (2004-2008), and consulting fees (2006-2008) from GlaxoSmithKline. EKS received an honorarium from Wyeth for a talk on COPD genetics in 2004. EKS received an honorarium from Bayer for a symposium at the ERS Meeting in 2005. EKS received honoraria for talks in 2007 and 2008 and consulting fees in 2008 from AstraZeneca.
PubMed Abstract | Publisher Full Text | Free Full Text

17. Shrine N, Guyatt AL, Erzurumluoglu AM, *et al.*: **New genetic signals for lung function highlight pathways and pleiotropy, and chronic obstructive pulmonary disease associations across multiple ancestries.** *bioRxiv.* 2018.
Publisher Full Text

18. Weedon MN, Lango H, Lindgren CM, *et al.*: **Genome-wide association analysis identifies 20 loci that influence adult height.** *Nat Genet.* 2008; **40**(5): 575–83.
PubMed Abstract | Publisher Full Text | Free Full Text

19. Zhou X, Baron RM, Hardin M, *et al.*: **Identification of a chronic obstructive pulmonary disease genetic determinant that regulates HHIP.** *Hum Mol Genet.* 2012; **21**(6): 1325–35.
PubMed Abstract | Publisher Full Text | Free Full Text

20. Collins SA, Lucas JS, Inskip HM, *et al.*: **HHIP, HDAC4, NCR3 and RARB polymorphisms affect fetal, childhood and adult lung function.** *Eur Respir J.* 2013; **41**(3): 756–7.
PubMed Abstract | Publisher Full Text | Free Full Text

21. Mahmoud O, Granell R, Tilling K, *et al.*: **Association of Height Growth in Puberty with Lung Function. A Longitudinal Study.** *Am J Respir Crit Care Med.* 2018; **198**(12): 1539–1548.
PubMed Abstract | Publisher Full Text | Free Full Text

22. 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.* 2012; **40**(Database issue): D930–4.
PubMed Abstract | Publisher Full Text | Free Full Text

23. Heemers HV, Verhoeven G, Swinnen JV: **Androgen activation of the sterol regulatory element-binding protein pathway: Current insights.** *Mol Endocrinol.* 2006; **20**(10): 2265–77.
PubMed Abstract | Publisher Full Text

24. Leimgruber C, Quintar AA, Peinetti N, *et al.*: **Testosterone Rescues the De-Differentiation of Smooth Muscle Cells Through Serum Response Factor/Myocardin.** *J Cell Physiol.* 2017; **232**(10): 2806–17.
PubMed Abstract | Publisher Full Text

25. 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**(3): 494–505.
PubMed Abstract | Publisher Full Text | Free Full Text

26. Kugler MC, Joyner AL, Loomis CA, *et al.*: **Sonic hedgehog signaling in the lung. From development to disease.** *Am J Respir Cell Mol Biol.* 2015; **52**(1): 1–13.
PubMed Abstract | Publisher Full Text | Free Full Text

27. Petrova R, Joyner AL: **Roles for Hedgehog signaling in adult organ homeostasis and repair.** *Development.* 2014; **141**(18): 3445–57.
PubMed Abstract | Publisher Full Text | Free Full Text

28. Lange P, Celli B, Agustí A, *et al.*: **Lung-Function Trajectories Leading to Chronic Obstructive Pulmonary Disease.** *N Engl J Med.* 2015; **373**(2): 111–22.
PubMed Abstract | Publisher Full Text

29. Bycroft C, Freeman C, Petkova D, *et al.*: **Genome-wide genetic data on ~500,000 UK Biobank participants.** *bioRxiv.* 2017.
Publisher Full Text

30. Battram T, Hoskins L, Hughes DA, *et al.*: **Coronary artery disease, genetic risk and the metabolome in young individuals [version 2; peer review: 2 approved].** *Wellcome Open Res.* 2018; **3**: 114.
PubMed Abstract | Publisher Full Text | Free Full Text

31. 1000 Genomes Project Consortium, Abecasis GR, Altshuler D, *et al.*: **A map of human genome variation from population-scale sequencing.** *Nature.* 2010; **467**(7319): 1061–73.
PubMed Abstract | Publisher Full Text | Free Full Text

32. Yang J, Ferreira T, Morris AP, *et al.*: **Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits.** *Nat Genet.* 2012; **44**(4): 369–75, S1-3.
PubMed Abstract | Publisher Full Text | Free Full Text

33. Yang J, Lee SH, Goddard ME, *et al.*: **GCTA: a tool for genome-wide complex trait analysis.** *Am J Hum Genet.* 2011; **88**(1): 76–82.
PubMed Abstract | Publisher Full Text | Free Full Text

34. Minelli C, van der Plaat DA, Leynaert B, *et al.*: **Age at puberty and risk of asthma: A Mendelian randomisation study.** *PLoS Med.* 2018; **15**(8): e1002634. following competing interests: CM and GDS are members of the Editorial Board of PLOS Medicine.
PubMed Abstract | Publisher Full Text | Free Full Text

35. Hao K, Bossé Y, Nickle DC, *et al.*: **Lung eQTLs to help reveal the molecular underpinnings of asthma.** *PLoS Genet.* 2012; **8**(11): e1003029.
PubMed Abstract | Publisher Full Text | Free Full Text

36. 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.
PubMed Abstract | Publisher Full Text | Free Full Text

37. Obeidat M, Miller S, Probert K, *et al.*: **GSTCD and INTS12 regulation and expression in the human lung.** *PLoS One.* 2013; **8**(9): e74630.
PubMed Abstract | Publisher Full Text | Free Full Text

38. FFawcett K, Obeidat M, Melbourne C, *et al.*: **Genome-wide sex interaction study summary statistics for lung function traits in UK Biobank.** *figshare.* Journal contribution. 2020.
http://www.doi.org/10.6084/m9.figshare.12298736.v1

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 1

Reviewer Report 05 March 2021

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

✔ **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.

Reviewer Report 21 July 2020

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

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.