

# A Comprehensive Evaluation of Potential Lung Function Associated Genes in the SpiroMeta General Population Sample

Ma'en Obeidat<sup>1</sup>, Louise V. Wain<sup>2</sup>, Nick Shrine<sup>2</sup>, Noor Kalsheker<sup>3</sup>, Maria Soler Artigas<sup>2</sup>, Emmanouela Repapi<sup>2,4</sup>, Paul R. Burton<sup>2</sup>, Toby Johnson<sup>5</sup>, Adaikalavan Ramasamy<sup>6,7</sup>, Jing Hua Zhao<sup>8</sup>, Guangju Zhai<sup>9</sup>, Jennifer E. Huffman<sup>10</sup>, Veronique Vitart<sup>10</sup>, Eva Albrecht<sup>11</sup>, Wilmar Igl<sup>12</sup>, Anna-Liisa Hartikainen<sup>13</sup>, Anneli Pouta<sup>14</sup>, Gemma Cadby<sup>15,16</sup>, Jennie Hui<sup>17,18,19</sup>, Lyle J. Palmer<sup>15,16</sup>, David Hadley<sup>20,21</sup>, Wendy L. McArdle<sup>22</sup>, Alicja R. Rudnicka<sup>20</sup>, Inês Barroso<sup>23,24</sup>, Ruth J. F. Loos<sup>8</sup>, Nicholas J. Wareham<sup>8</sup>, Massimo Mangino<sup>9</sup>, Nicole Soranzo<sup>9,23</sup>, Tim D. Spector<sup>9</sup>, Sven Gläser<sup>25</sup>, Georg Homuth<sup>26</sup>, Henry Völzke<sup>27</sup>, Panos Deloukas<sup>23</sup>, Raquel Granell<sup>28</sup>, John Henderson<sup>28</sup>, Ivica Grkovic<sup>29</sup>, Stipan Jankovic<sup>29</sup>, Lina Zgaga<sup>30</sup>, Ozren Polašek<sup>31</sup>, Igor Rudan<sup>32,29</sup>, Alan F. Wright<sup>10</sup>, Harry Campbell<sup>32</sup>, Sarah H. Wild<sup>32</sup>, James F. Wilson<sup>32</sup>, Joachim Heinrich<sup>33</sup>, Medea Imboden<sup>34</sup>, Nicole M. Probst-Hensch<sup>34,35</sup>, Ulf Gyllensten<sup>36</sup>, Åsa Johansson<sup>36</sup>, Ghazal Zaboli<sup>12</sup>, Linda Mustelin<sup>37</sup>, Taina Rantanen<sup>38</sup>, Ida Surakka<sup>39,40</sup>, Jaakko Kaprio<sup>37,39,40</sup>, Marjo-Riitta Jarvelin<sup>7,41,42,14</sup>, Caroline Hayward<sup>10</sup>, David M. Evans<sup>43</sup>, Beate Koch<sup>25</sup>, Arthur William Musk<sup>18,44,45</sup>, Paul Elliott<sup>46,7</sup>, David P. Strachan<sup>20</sup>, Martin D. Tobin<sup>2\*</sup>, Ian Sayers<sup>1</sup>, Ian P. Hall<sup>1\*</sup>, SpiroMeta Consortium<sup>47</sup>

1 Nottingham Respiratory Biomedical Research Unit, Division of Therapeutics and Molecular Medicine, University Hospital of Nottingham, Nottingham, United Kingdom, 2 Departments of Health Sciences and Genetics, University of Leicester, Leicester, United Kingdom, 3 School of Molecular Medical Sciences and Centre for Genetics and Genomics, Queen's Medical Centre, University of Nottingham, Nottingham, United Kingdom, 4 Ludwig Institute for Cancer Research, University of Oxford, Oxford, United Kingdom, 5 Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary, University of London, London, United Kingdom, 6 Respiratory Epidemiology and Public Health Group, National Heart and Lung Institute, Imperial College London, London, United Kingdom, 7 Department of Epidemiology and Biostatistics, Imperial College London, London, United Kingdom, 8 MRC Epidemiology Unit, Institute of Metabolic Science, Cambridge, United Kingdom, 9 Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom, 10 MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, Scotland, United Kingdom, 11 Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany, 12 Rudbeck Laboratory, Department of Genetics and Pathology, Uppsala University, Uppsala, Sweden, 13 Department of Clinical Sciences, Obstetrics and Gynecology, Institute of Clinical Medicine, University of Oulu, Oulu, Finland, 14 Department of Life Course and Services, National Institute for Health and Welfare, Oulu, Finland, 15 Ontario Institute for Cancer Research, Toronto, Canada, 16 Samuel Lunenfeld Research Institute, Toronto, Canada, 17 Molecular Genetics, PathWest Laboratory Medicine WA, Nedlands, Western Australia, Australia, 18 Busselton Population Medical Research Foundation, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia, 19 Schools of Population Health and Pathology and Laboratory Medicine, University of Western Australia, Crawley, Australia, 20 Division of Community Health Sciences, St George's University of London, London, United Kingdom, 21 Pediatric Epidemiology Center, University of South Florida, Tampa, Florida, United States of America, 22 ALSPAC Laboratory, School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom, 23 Wellcome Trust Sanger Institute, Cambridge, United Kingdom, 24 University of Cambridge Metabolic Research Labs, Institute of Metabolic Science Addenbrooke's Hospital Cambridge, Cambridge, United Kingdom, 25 Department of Internal Medicine B - Cardiology, Intensive Care, Pulmonary Medicine and Infectious Diseases, University of Greifswald, Greifswald, Germany, 26 Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald, Greifswald, Germany, 27 Institute for Community Medicine, SHIP/Clinical-Epidemiological Research, University of Greifswald, Greifswald, Germany, 28 School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom, 29 Croatian Centre for Global Health, The University of Split Medical School, Split, Croatia, 30 Andrija Stampar School of Public Health, Faculty of Medicine, University of Zagreb, Zagreb, Croatia, 31 Department of Public Health, University of Split, Split, Croatia, 32 Centre for Population Health Sciences, University of Edinburgh, Edinburgh, Scotland, United Kingdom, 33 Institute of Epidemiology I, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany, 34 University of Basel, Basel, Switzerland, 35 Swiss Tropical and Public Health Institute, Basel, Switzerland, 36 Rudbeck Laboratory, Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden, 37 Department of Public Health, University of Helsinki, Helsinki, Finland, 38 Department of Health Sciences and Gerontology Research Centre, University of Jyväskylä, Jyväskylä, Finland, 39 Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland, 40 National Institute for Health and Welfare, Helsinki, Finland, 41 Institute of Health Sciences, University of Oulu, Oulu, Finland, 42 Biocenter Oulu, University of Oulu, Oulu, Finland, 43 MRC Centre for Causal Analyses in Translational Epidemiology, School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom, 44 Department of Respiratory Medicine, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia, 45 Schools of Population Health and Medicine and Pharmacology, University of Western Australia, Crawley, Australia, 46 MRC-HPA Centre for Environment and Health, Imperial College London, London, United Kingdom, 47 SpiroMeta Consortium, Nottingham, Leicester, United Kingdom

## Abstract

**Rationale:** Lung function measures are heritable traits that predict population morbidity and mortality and are essential for the diagnosis of chronic obstructive pulmonary disease (COPD). Variations in many genes have been reported to affect these traits, but attempts at replication have provided conflicting results. Recently, we undertook a meta-analysis of Genome Wide Association Study (GWAS) results for lung function measures in 20,288 individuals from the general population (the SpiroMeta consortium).

**Objectives:** To comprehensively analyse previously reported genetic associations with lung function measures, and to investigate whether single nucleotide polymorphisms (SNPs) in these genomic regions are associated with lung function in a large population sample.

**Methods:** We analysed association for SNPs tagging 130 genes and 48 intergenic regions (+/−10 kb), after conducting a systematic review of the literature in the PubMed database for genetic association studies reporting lung function associations.

**Results:** The analysis included 16,936 genotyped and imputed SNPs. No loci showed overall significant association for FEV<sub>1</sub> or FEV<sub>1</sub>/FVC traits using a carefully defined significance threshold of  $1.3 \times 10^{-5}$ . The most significant loci associated with FEV<sub>1</sub> include SNPs tagging *MACROD2* ( $P = 6.81 \times 10^{-5}$ ), *CNTN5* ( $P = 4.37 \times 10^{-4}$ ), and *TRPV4* ( $P = 1.58 \times 10^{-3}$ ). Among ever-smokers, *SERPINA1* showed the most significant association with FEV<sub>1</sub> ( $P = 8.41 \times 10^{-5}$ ), followed by *PDE4D* ( $P = 1.22 \times 10^{-4}$ ). The strongest association with FEV<sub>1</sub>/FVC ratio was observed with *ABCC1* ( $P = 4.38 \times 10^{-4}$ ), and *ESR1* ( $P = 5.42 \times 10^{-4}$ ) among ever-smokers.

**Conclusions:** Polymorphisms spanning previously associated lung function genes did not show strong evidence for association with lung function measures in the SpiroMeta consortium population. Common *SERPINA1* polymorphisms may affect FEV<sub>1</sub> among smokers in the general population.

**Citation:** Obeidat M, Wain LV, Shrine N, Kalsheker N, Artigas MS, et al. (2011) A Comprehensive Evaluation of Potential Lung Function Associated Genes in the SpiroMeta General Population Sample. PLoS ONE 6(5): e19382. doi:10.1371/journal.pone.0019382

**Editor:** Malcolm Gracie Semple, University of Liverpool, United Kingdom

**Received:** February 11, 2011; **Accepted:** March 28, 2011; **Published:** May 20, 2011

**Copyright:** © 2011 Obeidat et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Cohort funding: ALSPAC: The UK Medical Research Council (Grant number: G990146), the Wellcome Trust and the University of Bristol provide core support for ALSPAC. B58C - WTCCC: The British 1958 Birth Cohort DNA collection was funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. Genotyping for the Wellcome Trust Case Control Consortium was funded by the Wellcome Trust grant 076113/B/04/Z. B58C - T1DGC: This research utilizes resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. T1DGC GWAS data were deposited by the Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research (CIMR), University of Cambridge (John Todd, Helen Stevens and Neil Walker), which is funded by Juvenile Diabetes Research Foundation International, the Wellcome Trust and the National Institute for Health Research Cambridge Biomedical Research Centre; the CIMR is in receipt of a Wellcome Trust Strategic Award (079895). EPIC: The EPIC Norfolk Study is funded by Cancer Research United Kingdom and the Medical Research Council. I.B. received funding from the Wellcome Trust (077016/Z/05/Z) and from the United Kingdom NIHR Cambridge Biomedical Research Centre. FTC: Academy of Finland Centre of Excellence in Complex Disease Genetics. Finnish Twin Study on Aging was funded by Academy of Finland and Finnish Ministry of Education. ENGAGE project grant agreement HEALTH-F4-2007-201413 and the European Union FP-5 GenomEUtwin Project (QLG2-CT-2002-01254). KORA S3: The KORA Augsburg studies were financed by the Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany and supported by grants from the German Federal Ministry of Education and Research (BMBF) in the context of the German National Genome Research Network (NGFN-2 and NGFN-plus). The authors research was supported within the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ. Genotyping for the KORA F3/500K Study which included the lung function characterized KORA C sample was partly funded by UBS Wealth Foundation Grant BA29s8Q7-DZZ. Korcula: The Korcula study (Croatia) was funded by grants from the Medical Research Council (UK) and Republic of Croatia Ministry of Science, Education and Sports research grants to I.R. (Vis) (108-1080315-0302). NFBC1966: Financial support was received from the Academy of Finland (project grants 104781, 1114194, 120315 and Center of Excellence in Complex Disease Genetics), Oulu University Hospital, Biocenter Oulu, University of Oulu, Finland, the European Commission (EURO-BLCS, Framework 5 award QLG1-CT-2000-01643), NHLBI grant 5R01HL087679-02 through the STAMPEED program (1RL1MH083268-01), ENGAGE project and grant agreement HEALTH-F4-2007-201413, and the Medical Research Council (studentship grant G0500539). The DNA extractions, sample quality controls, biobank up-keeping and aliquotting was performed in the National Institute for Health and Welfare, Biomedicum Helsinki, Finland and supported financially by the Academy of Finland and Biocentrum Helsinki. A.R. was supported by the European Commission as part of GABRIEL (a multidisciplinary study to identify the genetic and environmental causes of asthma in the European Community) contract number 018996 under the Integrated Program LSH-2004-1.2.5-1 Post genomic approaches to understand the molecular basis of asthma aiming at a preventive or therapeutic control. NSPHS: EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP grant number 01947 (LSHG-CT-2006-01947). The Northern Swedish Population Health Study (NSPHS) was funded by the Swedish Medical Research Council (project number K2007-66X-20270-01-3). The computations were performed on UPPMAX (<http://www.uppmx.uu.se>) resources under Project p2008027. ORCADES: The ORCADES study was funded by the Chief Scientist Office of the Scottish Government, the Royal Society and the MRC Human Genetics Unit. DNA extraction was performed at the Wellcome Trust Clinical Research Facility in Edinburgh. Genotyping was funded by the European Union Framework Programme 6 EUROSPAN project. SHIP: SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the German Asthma and COPD Network (COSYCONET; BMBF grant 01GI0883), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg-West Pomerania. The University of Greifswald is a member of the 'Center of Knowledge Interchange' program of the Siemens AG. TwinsUK: TwinsUK is funded by the Wellcome Trust; the Arthritis Research Campaign; European Community's Seventh Framework Programme (FP7/2007–2013)/grant agreement HEALTH-F2-2008-201865-GEFOS and Seventh Framework Programme (FP7/2007–2013). ENGAGE project grant agreement HEALTH-F4-2007-201413 and the European Union FP-5 GenomEUtwin Project (QLG2-CT-2002-01254). The study also receives support from the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London. TDS is an NIHR senior Investigator. The project also received support from a Biotechnology and Biological Sciences Research Council (BBSRC) project grant. Vis: The Vis study (Croatia) was funded by grants from the Medical Research Council (UK), European Commission Framework 6 project EUROSPAN (Contract No. LSHG-CT-2006-018947) and Republic of Croatia Ministry of Science, Education and Sports research grants to I.R. (108-1080315-0302). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: lan.hall@nottingham.ac.uk (IH); mt47@le.ac.uk (MT)

## Introduction

Pulmonary function is usually assessed by measurement of forced expiratory volume in one second (FEV<sub>1</sub>), forced vital capacity (FVC), and the ratio of FEV<sub>1</sub> to FVC. The measurements are integral to the diagnosis of chronic obstructive pulmonary disease (COPD), and also are important long term predictors of population morbidity and mortality [1]. Reduced FEV<sub>1</sub>/FVC defines airways obstruction; whereas reduced FEV<sub>1</sub> grades the severity of obstruction [2].

Pulmonary function is determined by both environmental and genetic factors. Tobacco smoking is the major environmental risk factor for the development of COPD. A genetic contribution to pulmonary function is well established with heritability estimates reaching 77 percent for FEV<sub>1</sub> [3]. Linkage analyses within families have previously identified multiple genomic regions associated with spirometry measures and respiratory diseases. In addition, candidate gene studies have identified more than 100 genes which have been suggested to contribute to variability in lung function. The majority have been studied because of their potential pathophysiological role in the development of COPD. Some genes have been examined for association with lung function measurements in individuals with other specific respiratory diseases (most commonly asthma), or to a lesser extent, in the general population. With the exception of *SERPINA1*, which is the best documented genetic risk factor to influence the development of COPD [4], these genes have not shown consistent associations across different studies [5,6].

Recently, we established the SpiroMeta consortium and published a large collective meta-analysis of lung function genome-wide association studies (GWAS) in 20,288 individuals of European origin, with follow-up of top SNPs in a further 54,276 individuals [7,8]. Our study confirmed the hedgehog interacting protein (*HHIP*) association previously published [9,10] and identified five new loci associated with FEV<sub>1</sub> or FEV<sub>1</sub>/FVC ratio including tensin 1 gene (*TNSI*), glutathione S-transferase, C-terminal domain containing (*GSTCD*), 5-hydroxytryptamine receptor 4 (*HTR4*), advanced glycosylation end product-specific receptor (*AGER*), and thrombospondin, type I, domain containing 4 (*THSD4*). A study with similar design by the CHARGE consortium also identified *HHIP*, *AGER*, *HTR4*, and *GSTCD*, and in addition suggested a potential role of five additional genes (G protein-coupled receptor 126 (*GPR126*), a disintegrin and metalloproteinase domain 19 (*ADAM19*), family with sequence similarity 13, member A (*FAM13A*), patched homolog 1 (*PTCH1*), and phosphotyrosine interaction domain containing (*PIDI*) [8].

The identification of these genes offers potential insight into the pathophysiology of altered lung function. The SpiroMeta consortium provides a powerful resource in which to study genetic associations with lung function. We aimed to comprehensively evaluate whether genes studied in candidate gene or small genome-wide association studies, and reported to be associated with lung function or COPD in these studies, were associated with lung function measures in this large general population sample.

## Results

### Literature search

The literature search identified 1719 publications. Of these, 104 reported one or more genetic associations: these are listed in text S1 in the online supporting information. These publications varied according to their study designs and the populations studied. 47 papers reported association with COPD using case control or

family based designs. The remaining literature reported association with lung function traits within populations with specific respiratory diseases (asthma (26) and COPD (17)), or in general population cohorts (14). Nine publications studied other populations which included patients with cystic fibrosis (2), *SERPINA1* deficiency (2), cotton and grain workers (2), lung cancer (1), fire fighters (1) and post myocardial infarction (MI) patients (1). Some papers reported more than one endpoint.

These 104 relevant publications identified 130 genes and 48 intergenic SNPs. We investigated association between FEV<sub>1</sub> and FEV<sub>1</sub>/FVC and each of the 16,936 genotyped and imputed SNPs spanning these regions in the SpiroMeta dataset.

### Contribution of all tested genes to lung function measures in SpiroMeta

Quantile-quantile (Q-Q) plots did not show large deviations between observed and expected P values for FEV<sub>1</sub> and FEV<sub>1</sub>/FVC in all participants and for FEV<sub>1</sub>/FVC in ever-smokers (Figure 1). The plot of FEV<sub>1</sub> in ever-smokers, however, shows slight deviations for high signal SNPs. The genomic inflation factor,  $\lambda$  for FEV<sub>1</sub> is 0.83 in all individuals and 1.05 in smokers;  $\lambda$  for FEV<sub>1</sub>/FVC in all individuals is 0.92 and 1.13 in smokers.

Using the Bonferroni corrected P value threshold of  $1.3 \times 10^{-5}$ , none of the tested SNPs demonstrated significant association with either FEV<sub>1</sub> or FEV<sub>1</sub>/FVC.

### Association results in all individuals

In order to examine possible signals in greater detail, we also explored region plots for the top SNPs identified (SNPs with the lowest P values). The three top loci with the most significant P values for all regions tested in all individuals are presented in table 1.

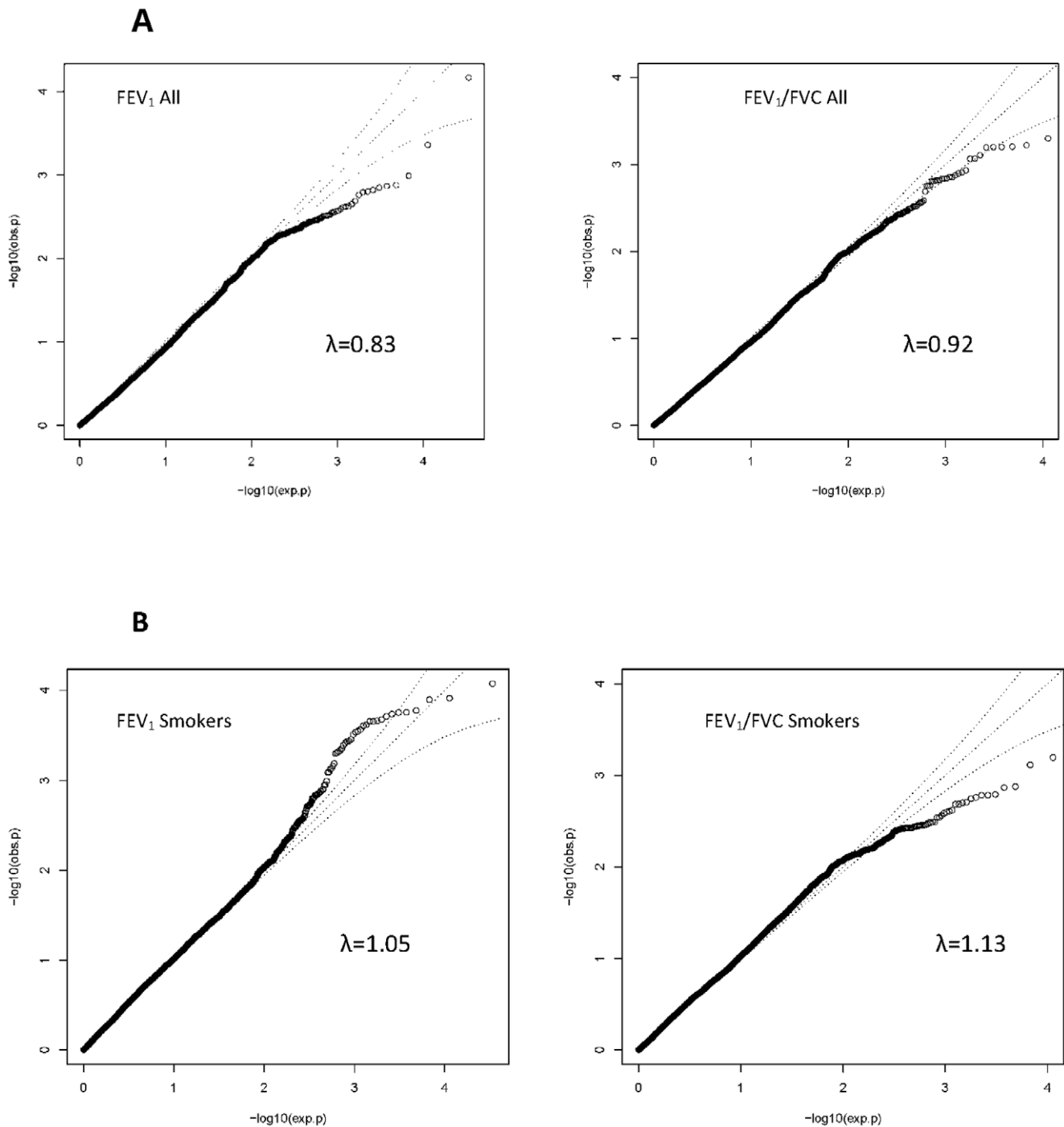
Among all individuals, the strongest association with FEV<sub>1</sub> was with rs204652 in MACRO domain containing 2 (*MACROD2*) on chromosome 20. SNP rs17133553 in Contactin 5 (*CNTN5*) on chromosome 11 was the second top locus for association with FEV<sub>1</sub> and third for FEV<sub>1</sub>/FVC ratio. SNP rs803450 in Methylene tetrahydrofolate dehydrogenase (NADP+ dependent) 1-like (*MTHFD1L*) on chromosome 6 showed association with FEV<sub>1</sub> in all individuals. For FEV<sub>1</sub>/FVC ratio in all individuals the strongest association was with rs3887893 in ATP-binding cassette, sub-family C, member 1 (*ABCC1*) on chromosome 16, the second strongest signal was for rs11155818 in estrogen receptor 1 (*ESR1*) on chromosome 6.

The region association plots around the most significant SNPs associated with FEV<sub>1</sub> and FEV<sub>1</sub>/FVC in all individuals provide little evidence from supporting SNPs to suggest strong regions of association in *MACROD2*, *CNTN5*, *MTHFD1L*, and *ESR1*, and *ABCC1* in these data (See figure S1 in the online supporting information).

### Association results in ever-smokers

To study the impact of smoking on potential genetic associations with lung function, we repeated the analysis restricted to individuals who had ever smoked (ever-smokers). The most significant loci identified are shown in table 1.

Among ever-smokers, rs3748312 in serpin peptidase inhibitor, clade A, member 1 (*SERPINA1*) on chromosome 14, also known as alpha-1 antitrypsin (*AAT*) showed the strongest association with FEV<sub>1</sub>. SNP rs298028 in Phosphodiesterase 4D, cAMP-specific (phosphodiesterase E3 duncce homolog, Drosophila) (*PDE4D*) on chromosome 5 showed the second strongest association with FEV<sub>1</sub>, followed by *MACROD2*.



**Figure 1. Quantile-quantile (Q-Q) plots of association results for FEV<sub>1</sub> and FEV<sub>1</sub>/FVC.** The Q-Q plot of association is shown for all individuals in Panel A, and separately for ever-smokers only in panel B. Q-Q plots compare the observed P values (obs.p) to expected P values (exp.p) on the logarithmic scale under the null hypothesis of no significant association.  $\lambda$ : Lambda.  
doi:10.1371/journal.pone.0019382.g001

The strongest association with FEV<sub>1</sub>/FVC ratio among smokers was observed with rs9322335 in 1(*ESR1*) on chromosome 6. The second strongest association was rs1864271 in with rhomboid domain containing 1 (*RHBDD1*) on chromosome 2, followed by rs1738567 in (*MTHFD1L*) on chromosome 6. The region association plots for *SERPINA1* and *PDE4D* among ever-smokers (figure 2) show some supportive evidence for the association of these two loci. The region association

plots for the additional loci among ever-smokers reported in table 1 are shown in figure S2 in the online supporting information.

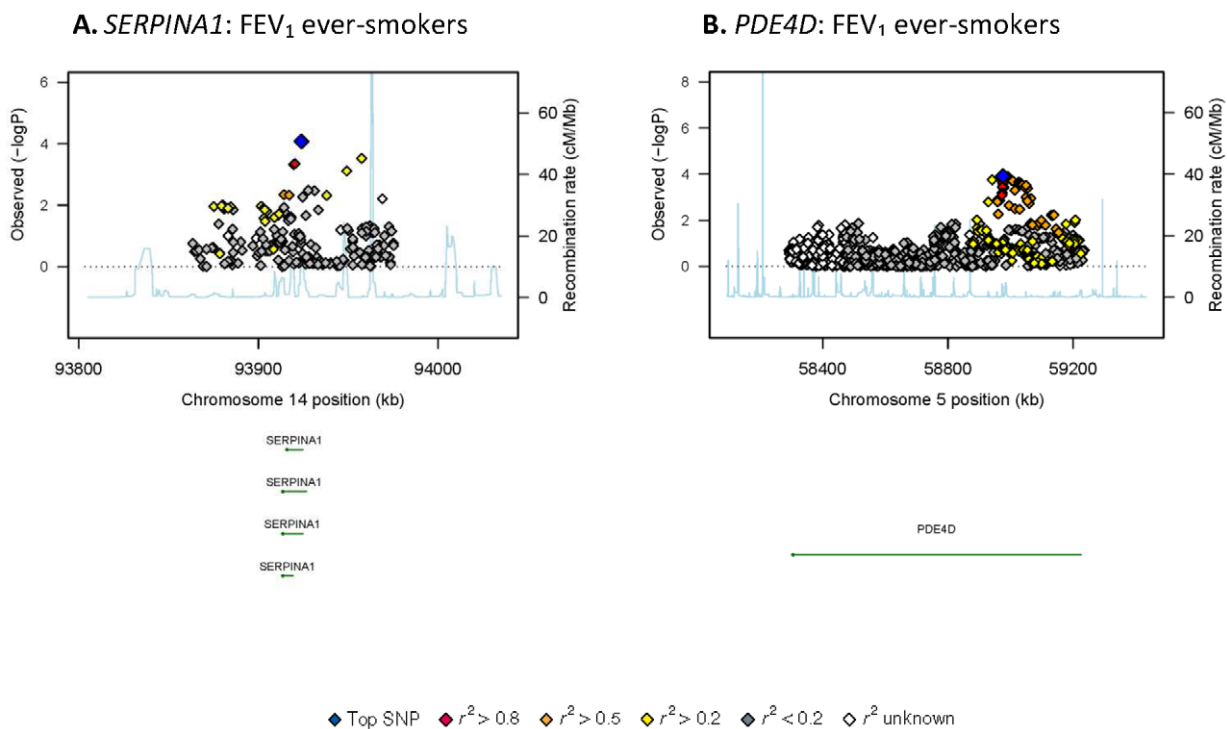
#### Association results excluding loci identified in previous GWAS

Because some of the regions identified were observed in the previously published small GWAS studies included in our

**Table 1.** Association results for the three most significantly associated loci.

Gene	Locus	SNP	SNP function	coded allele	Coded allele frequency	N eff	Beta	Se	P
<b>FEV<sub>1</sub> All individuals</b>									
<i>MACROD2</i>	20p12.1	rs204652	Intron	G	0.983	13551	-0.187	0.047	$6.81 \times 10^{-5}$
<i>CNTN5</i>	11q21-q22.2	rs17133553	Intron	T	0.966	13669	-0.100	0.029	$4.37 \times 10^{-4}$
<i>MTHFD1L</i>	6q25.1	rs803450	Intron	G	0.625	18497	0.036	0.011	$1.03 \times 10^{-3}$
<b>FEV<sub>1</sub> Smokers</b>									
<i>SERPINA1</i>	14q32.13	rs3748312	Intron	T	0.167	9338	0.085	0.022	$8.41 \times 10^{-5}$
<i>PDE4D</i>	5q12	rs298028	Intron	T	0.283	10829	-0.069	0.018	$1.22 \times 10^{-4}$
<i>MACROD2</i>	20p12.1	rs204652	Intron	G	0.983	6872	-0.251	0.067	$1.67 \times 10^{-4}$
<b>FEV<sub>1</sub>/FVC All individuals</b>									
<i>ABCC1</i>	16p13.1	rs3887893	Intron	T	0.625	15509	-0.043	0.012	$4.38 \times 10^{-4}$
<i>ESR1</i>	6q25.1	rs11155818	Intron	G	0.992	12571	0.185	0.053	$5.02 \times 10^{-4}$
<i>CNTN5</i>	11q21-q22.2	rs1216170	Intron	C	0.284	18863	0.039	0.011	$6.02 \times 10^{-4}$
<b>FEV<sub>1</sub>/FVC Smokers</b>									
<i>ESR1</i>	6q25.1	rs9322335	Intron	T	0.217	9495	-0.061	0.018	$5.42 \times 10^{-4}$
<i>RHBDD1</i>	2q36.3	rs1864271	Intron	G	0.708	7385	0.066	0.019	$6.40 \times 10^{-4}$
<i>MTHFD1L</i>	6q25.1	rs1738567	Intron	C	0.367	10668	0.046	0.014	$1.33 \times 10^{-3}$

The table shows the three most significant loci associated with FEV<sub>1</sub> and FEV<sub>1</sub>/FVC ratio in all individuals and ever-smokers. N eff: the effective sample size. Beta: regression coefficient on a transformed scale. Se: standard error. P: P value. Coded allele frequency based on HapMap Release 24. *MACROD2*: MACRO domain containing 2. *CNTN5*: Contactin 5. *MTHFD1L*: Methylene tetrahydrofolate dehydrogenase (NADP+ dependent) 1-like. *SERPINA1*: serpin peptidase inhibitor, clade A, member 1 (alpha-1 antitrypsin AAT). *PDE4D*: Phosphodiesterase 4D, cAMP-specific (phosphodiesterase E3 duncce homolog, Drosophila). *ABCC1*: ATP-binding cassette, sub-family C, member 1. *ESR1*: estrogen receptor 1. *RHBDD1*: rhomboid domain containing 1. doi:10.1371/journal.pone.0019382.t001



**Figure 2. Regional association plots for *SERPINA1* (A) and *PDE4D* among ever-smokers (B) in SpiroMeta.** Statistical significance of each SNP on the  $-\log_{10}$  scale as a function of chromosome position (NCBI build 36). The sentinel SNP at each locus is shown in blue; the correlations ( $r^2$ ) of each of the surrounding SNPs to the sentinel SNP are shown in the indicated colours. The relevant trait (FEV<sub>1</sub> or FEV<sub>1</sub>/FVC ratio) is indicated for each plot. Recombination rate is shown in pale blue. doi:10.1371/journal.pone.0019382.g002

literature search, we also present the top three genes for the relevant end points after excluding GWAS hits (Table 2).

The additional genes identified in this analysis for association with FEV<sub>1</sub> among all individuals were the transient receptor potential cation channel, subfamily V, member 4 (*TRPV4*) on chromosome 12, and N-acetyltransferase 2 (*NAT2*) on chromosome 8. Among ever-smokers, association results for FEV<sub>1</sub> identified B-cell CLL/lymphoma 2 (*BCL2*) on chromosome 18. Association results for FEV<sub>1</sub>/FVC ratio identified allograft inflammatory factor 1 (*AIFI*) on chromosome 6 among all individuals, and cluster of differentiation; CD22 molecule (*CD22*) on chromosome 19 among ever-smokers. The region association plots for the most significant loci in table 2 and not presented earlier are shown in figure S3 in the online supporting information. The plots show some additional support for all presented loci except for *ABCC1* among ever-smokers.

## Discussion

In the SpiroMeta study, we generated a comprehensive dataset to analyse associations between genetic variants and lung function in the general population [7]. There have been many small previous studies, mostly of individual candidate genes examining association with lung function, which have produced conflicting results. Therefore, in this paper, we undertook a comprehensive literature review to identify relevant gene regions and analysed potential associations with FEV<sub>1</sub> and FEV<sub>1</sub>/FVC ratio in all individuals within SpiroMeta. In addition, given the impact of smoking on lung function, we also analysed the associations separately in ever-smokers. There were no strong association signals in never-smokers group (data available on request).

The main conclusion from this study is that, within 178 previously reported regions, we found no SNP associations which exceeded the significance threshold ( $P < 1.3 \times 10^{-5}$ ) we employed after correction for multiple testing. Our results suggest these regions do not constitute major genetic determinants of lung function measures at the general population level. The lack of replication and sometimes contradicting results in previous studies may reflect the fact that many previously reported associations came from studies with small sample sizes, possibly leading to false positive results.

Despite the failure to identify any overall significant contribution of a single SNP from previously reported genes to lung function, there are some potentially interesting signals apparent from the region plots suggesting that there may be a small signal from variants in some of the genes of interest.

*SERPINA1* showed the strongest association with FEV<sub>1</sub> among smokers ( $8.41 \times 10^{-5}$ ). It encodes alpha-1 Antitrypsin protein (AAT), mainly produced in the liver and has the primary role of inhibiting neutrophil elastase in the lungs [11]. Protein variants of this gene have been classified based on their migration in an isoelectric pH gradient from A to Z. Among Caucasians, the M allele is the most common allele with six subtypes: M1–M6 with allele frequencies greater than 95 percent and associated with normal AAT levels. The common deficiency variants; S (frequency 0.02–0.03) and Z (frequency 0.01–0.03), are associated with mild and severe reductions in serum AAT levels, respectively [11,12]. The  $r^2$  between the Z allele rs28929474 and rs3748312 is 0.08 (based on 1000 Genomes Project pilot 1 data from 120 CEU individuals). Our top SNP, rs3748312, is in LD ( $r^2 = 0.603$ ) with the M1 allele SNP rs6647, but is in very weak LD with M2 rs709932 ( $r^2 = 0.033$ ) and M3 rs1303 ( $r^2 = 0.051$ ). The S allele SNP rs45551939 (merged into rs17580) was not found in HapMap

**Table 2.** Association results for the most significant loci excluding genes identified in GWAS.

gene	locus	SNP	SNP function	coded allele	Coded allele Frequency	N eff	Beta	se	P
<b>FEV<sub>1</sub> All</b>									
<i>PDE4D</i>	5q12	rs172362	Intron	C	0.133	18929	-0.048	0.015	$1.36 \times 10^{-3}$
<i>TRPV4</i>	12q24.1	rs3742030	Missense	G	0.992	17591	-0.110	0.035	$1.58 \times 10^{-3}$
<i>NAT2</i>	8p22	rs6988857	Intergenic	T	0.241	20080	0.035	0.011	$2.37 \times 10^{-3}$
<b>FEV<sub>1</sub> Smokers</b>									
<i>SERPINA1</i>	14q32.13	rs3748312	Intron	T	0.167	9338	0.085	0.022	$8.41 \times 10^{-5}$
<i>PDE4D</i>	5q12	rs298028	Intron	T	0.283	20829	-0.069	0.018	$1.22 \times 10^{-4}$
<i>BCL2</i>	18q21.3	rs2850760	Intron	T	0.425	10888	0.050	0.014	$4.66 \times 10^{-4}$
<b>FEV<sub>1</sub>/FVC All</b>									
<i>ABCC1</i>	16p13.1	rs3887893	Intron	T	0.625	15509	-0.043	0.012	$4.38 \times 10^{-4}$
<i>ESR1</i>	6q25.1	rs11155818	Intron	G	0.992	12571	0.185	0.053	$5.02 \times 10^{-4}$
<i>AIFI</i>	6p21.3	rs3132451	5' near gene	G	0.883	19889	0.045	0.014	$7.84 \times 10^{-4}$
<b>FEV<sub>1</sub>/FVC Smokers</b>									
<i>ESR1</i>	6q25.1	rs9322335	Intron	T	0.217	9495	-0.061	0.018	$5.42 \times 10^{-4}$
<i>ABCC1</i>	16p13.1	rs3887893	Intron	T	0.625	8156	-0.054	0.017	$1.36 \times 10^{-3}$
<i>CD22</i>	19q13.1	rs7251526	Intron	T	0.246	8328	-0.054	0.018	$2.00 \times 10^{-3}$

The table shows association results for the three most significant loci associated with FEV<sub>1</sub> and FEV<sub>1</sub>/FVC, in all individuals and in ever-smokers after excluding genes identified in GWAS. N eff: the effective sample size. Beta: regression coefficient on a transformed scale. Se: standard error. P: P value. Coded allele frequency based on HapMap Release 24. *PDE4D*: Phosphodiesterase 4D, cAMP-specific (phosphodiesterase E3 dunce homolog, Drosophila). *TRPV4*: transient receptor potential cation channel, subfamily V, member 4. *NAT2*: N-acetyltransferase 2. *SERPINA1*: serpin peptidase inhibitor, clade A, member 1 (alpha-1 antitrypsin AAT). *BCL2*: B-cell CLL/lymphoma 2. *ABCC1*: ATP-binding cassette, sub-family C, member 1. *AIFI*: allograft inflammatory factor 1. *ESR1*: estrogen receptor 1. *CD22*: cluster of differentiation; CD22 molecule.

doi:10.1371/journal.pone.0019382.t002

(version24). It is possible that the signal observed in our data is due to variants with effects on gene expression and/or protein levels, and this idea is supported by a previous study showing novel variants in *SERPINA1* to be associated with increased susceptibility to COPD independently of the Z allele [13]. The relatively strong signal observed in our study suggests a possible role for variants in *SERPINA1* in smokers at the general population level beyond that observed in carriers of known deficient alleles.

The *PDE4D* gene encodes the type 4D phosphodiesterase, which degrades cyclic adenosine monophosphate (cAMP), an important signal transduction molecule in all cell types. Polymorphisms within *PDE4D* have been associated with stroke [14], and bone mineral density [15]. *PDE4D* is the most dominant phosphodiesterase in the lungs and plays an important role in regulating airway smooth muscle contractility [16] demonstrated by *PDE4D* knockout mice lacking response to methacholine [17]. A study in a Japanese population reported association of one *PDE4D* SNP (rs829259) and a haplotype consisting of rs10075508 and one interleukin 13 (*IL13*) SNP with COPD [18]. SNP rs829259 was not associated with FEV<sub>1</sub> in all individuals ( $P = 0.68$ ) and in smokers ( $P = 0.21$ ) in our study, and SNP rs10075508 was not genotyped or imputed in SpiroMeta. A recent GWAS has also identified *PDE4D* as an asthma susceptibility gene [19], however, none of the top 5 SNPs associated with asthma is present in our dataset, and the linkage disequilibrium (LD) with SNPs in SpiroMeta is low, so it is difficult to comment on their contribution to lung function measures in our study.

Our study has a number of strengths. First, we have power to detect associations of small magnitude, with data on 20,288 individuals from 14 European studies with more than 2.5 million genotyped and imputed SNPs. Second, we aimed to minimise Type 1 error whilst taking appropriate account of the correlation between neighbouring SNPs. Finally, the literature search was designed to be comprehensive to include all reported genetic variants with effect on lung function irrespective of disease status or ethnicity. To our knowledge, this is the first study to comprehensively evaluate the role of previously associated genes in a large genome-wide association study.

However, it is important to recognise the limitations of our study. We have tested for association in a general population sample; the magnitude of effect of these genetic variants may be greater in populations enriched with individuals with respiratory diseases such as asthma and COPD. Second, we have tested with cross sectional lung function measures. Some of the variants tested might affect longitudinal changes by accelerating or decelerating the decline in lung function, although this would still be expected to result in effects evident in cross sectional data. Third, the power of our study to detect associations of SNPs with modest effect sizes on lung function was limited given our relatively conservative approach to multiple testing, therefore we cannot rule out a real but modest effect of some of these loci on lung function and susceptibility to respiratory diseases in the general population. Alternative approaches could be to utilise a priori evidence about the reported direction of effect and a priori assumptions about the likely presence of multiple causal variants. Fourth, we tested for association with lung function measures among individuals of European ancestry, and the contribution of these variants to lung function in other populations may vary. Finally, the coverage of tested genetic regions varies depending on the genome-wide arrays used and imputation quality metrics.

In conclusion, we have shown that none of the SNPs tagging the genes previously reported to determine lung function were significantly associated with FEV<sub>1</sub> or FEV<sub>1</sub>/FVC ratio in the SpiroMeta general population study. We found some evidence to

suggest a possible contribution for the *SERPINA1* and *PDE4D* loci to lung function in smokers which warrant further study. As a resource to the scientific community we have provided the complete association results (Dataset S1) in the online supporting information.

## Methods

### Systematic Literature search

We conducted a literature search in PubMed in October 2009 for genetic association studies of lung function measurements and/or COPD. The search terms used were: “Lung function” OR “pulmonary function” OR “FEV<sub>1</sub>” OR “FEV<sub>1</sub>/FVC” OR “Forced Expiratory Volume” OR “Forced Vital Capacity” OR “ASTHMA” OR “COPD” OR “BRONCHIAL HYPERRESPONSIVENESS” OR “BHR” OR “obstructive lung disease” AND “SNP” OR “single nucleotide polymorphism” OR “polymorphism” OR “gene” OR “genetic” OR “genom\*” OR “variation” AND “Linkage” OR “association”

From the search results, we included relevant papers reporting only positive association results. For the three GWAS papers identified, we took a more inclusive approach and included all loci presented in the publication body, and not just those meeting genome-wide significance. We excluded papers reporting associations with respiratory diseases (e.g. asthma) without association with lung function measurements.

### Statistical analysis

The genes and intergenic SNPs identified in the relevant literature were evaluated in the SpiroMeta dataset using an extended region of  $\pm 10$  kilobases (kb) from the gene coordinates downloaded from the UCSC genome browser (we used the SNP coordinate  $\pm 10$  kb for intergenic SNPs). Meta-analysis association results for SNPs in these ( $\pm 10$  kb extended) regions were extracted from the SpiroMeta dataset for both FEV<sub>1</sub> and FEV<sub>1</sub>/FVC in all individuals and separately in ever-smokers. The complete cohort descriptions, study design and methods have been previously reported [7], but we provide here a brief summary. At study level, non-genotyped SNPs were imputed using standard approaches [18,20] to facilitate meta-analysis of studies employing different genotyping platforms. Thus up to 2,705,257 SNPs were tested for association with FEV<sub>1</sub> and FEV<sub>1</sub>/FVC using additive models and adjusting for age, sex, height and ancestry principal components. Then, the results were meta-analysed across studies using inverse variance weighting. Genomic control was applied at the study level and after the meta-analysis to correct for test inflation due to population stratification [21]. We excluded SNPs which were not well measured or imputed in the study (identifiable by an “effective sample size” of  $< 50\%$  of the total sample size) [7]. In all, we identified 16,936 genotyped and imputed SNPs in the gene and intergenic regions described above which met our inclusion criteria.

In order to correct for multiple testing of SNPs in linkage disequilibrium we used Li and Ji’s [22] method for calculating the effective number of independent tests from pairwise SNP correlations. Pairwise SNP correlations were obtained from reference genotypes of 1468 subjects in the Busselton study [23]. We estimated that the association tests for the 16,936 highly correlated SNPs we selected in the regions of interest equated to 3,891 independent tests.

To maintain a Type 1 error rate of 5%, we adjusted the significance threshold using a Bonferroni correction ( $0.05/3891$ ). Thus a threshold of  $1.3 \times 10^{-5}$  was used to determine statistical significance.

## Supporting Information

**Figure S1** Regional association plots of the most significant lung function-associated loci among all individuals in SpiroMeta (A–F). Statistical significance of each SNP on the  $-\log_{10}$  scale as a function of chromosome position (NCBI build 36). The sentinel SNP at each locus is shown in blue; the correlations ( $r^2$ ) of each of the surrounding SNPs to the sentinel SNP are shown in the indicated colours. The relevant trait (FEV<sub>1</sub> or FEV<sub>1</sub>/FVC ratio) is indicated for each plot. Recombination rate is shown in pale blue. (TIFF)

**Figure S2** Regional association plots of the most significant lung function-associated loci among ever-smokers in SpiroMeta (A–D). Statistical significance of each SNP on the  $-\log_{10}$  scale as a function of chromosome position (NCBI build 36). The sentinel SNP at each locus is shown in blue; the correlations ( $r^2$ ) of each of the surrounding SNPs to the sentinel SNP are shown in the indicated colours. The relevant trait (FEV<sub>1</sub> or FEV<sub>1</sub>/FVC ratio) is indicated for each plot. Recombination rate is shown in pale blue. (TIF)

**Figure S3** Regional association plots of the most significant lung function-associated loci (A–G) after excluding genes identified in GWAS. Statistical significance of each SNP on the  $-\log_{10}$  scale as a function of chromosome position (NCBI build 36). The sentinel SNP at each locus is shown in blue; the correlations ( $r^2$ ) of each of the surrounding SNPs to the sentinel SNP are shown in the indicated colours. The relevant trait (FEV<sub>1</sub> or FEV<sub>1</sub>/FVC ratio) and whether it is in all individuals or ever-smokers is indicated for each plot. Recombination rate is shown in pale blue. (TIF)

**Text S1** The 104 relevant publications identified in the literature search. (DOC)

**Dataset S1** Complete FEV<sub>1</sub> and FEV<sub>1</sub>/FVC association results for all individuals and separately for ever-smokers. (XLS)

## Acknowledgments

### ALSPAC

We thank the Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute for generating the ALSPAC GWA data.

### B58C – T1DGC

We acknowledge use of the DNA from the British 1958 Birth Cohort collection, funded by the Medical Research Council and Wellcome Trust. We thank the Avon Longitudinal Study of Parents and Children laboratory in Bristol and the British 1958 Birth Cohort team, including S. Ring, R. Jones, M. Pembrey, W. McArdle, D.P. Strachan and P. Burton for preparing and providing the control DNA samples.

### NFBC1966

We thank Professor Paula Rantakallio (launch of NFBC1966 and 1986), Ms Outi Tornwall and Ms Minttu Jussila (DNA biobanking).

### ORCADES

As a EUROSPAN partner, we thank Yurii Aulchenko, Department of Epidemiology, Erasmus University Medical Center and Anatoly V. Kirichenko, Institute of Cytology and Genetics, Siberian Division of Russian

Academy of Sciences, Novosibirsk, Russia for respectively performing imputation of the genotypic data and providing IT facilities. The ORCADES study would like to acknowledge the invaluable contributions of Lorraine Anderson (; lorraine.anderson@ed.ac.uk) and the research nurses in Orkney.

### Vis

The Vis study would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and Zagreb Medical Schools, Institute for Anthropological Research in Zagreb and Croatian Institute for Public Health.

### BHS

The Busselton Health Study acknowledges the generous support for the 1994/5 follow-up study from Healthway, Western Australia. The Busselton Health Study is supported by The Great Wine Estates of the Margaret River region of Western Australia. The BHS gratefully acknowledges the assistance of the Western Australian DNA Bank (NHMRC Enabling Facility) with DNA samples and the support provided by the Western Australian Genetic Epidemiology Resource (NHMRC Enabling Facility) for this study.

## Author Contributions

Conceived and designed the experiments: IPH I. Sayers LVW MO NK ER N. Shrine MSA PRB TJ MDT. Performed the experiments: MO IPH I. Sayers LVW MDT. Analyzed the data: LVW MO N. Shrine MSA I. Sayers IPH MDT. Contributed reagents/materials/analysis tools: LVW MO N. Shrine MSA I. Sayers IPH MDT. Wrote the paper: MO I. Sayers N. Shrine LVW MDT IPH. ALSPAC Project conception, design and management: J. Henderson RG. ALSPAC Phenotype collection and data management: J. Henderson RG. ALSPAC Genotyping: PD. ALSPAC Data analysis: DME. B58C - WTCCC Project conception, design and management: DPS. B58C - WTCCC Phenotype collection and data management: DPS ARR. B58C - WTCCC Genotyping: WLM. B58C - WTCCC Data analysis: ARR. B58C - T1DGC Data analysis: DPS DH. EPIC Project conception, design and management: IB RJFL NJW JHZ. EPIC Phenotype collection and data management: NJW. EPIC Genotyping: IB RJFL NJW JHZ. EPIC Data analysis: RJFL JHZ. FTC Project conception, design and management: JK TR. FTC Phenotype collection and data management: JK LM TR. FTC Genotyping: JK I. Surakka. FTC Data analysis: I. Surakka LM. KORA S3 Project conception, design and management: J. Heinrich. KORA S3 Phenotype collection and data management: J. Heinrich. KORA S3 Data analysis: EA MI NMP-H. Korcula Project conception, design and management: HC IG SJ IR AFW LZ. Korcula Phenotype collection and data management: IG SJ OP IR LZ. Korcula Data analysis: CH JEH VV. NFBC1966 Investigators: PE M-RJ AP AR A-LH. NFBC1966 Project conception, design and management: PE M-RJ A-LH AP. NFBC1966 Phenotype collection and data management: PE M-RJ AP. NFBC1966 Genotyping: PE M-RJ. NFBC1966 Data analysis: AR. NSPHS Project conception, design and management: UG. NSPHS Phenotype collection and data management: G. Zaboli. NSPHS Data analysis: WI AJ. ORCADES Project conception, design and management: HC SHW JFW AFW. ORCADES Phenotype collection and data management: HC SHW JFW. ORCADES Genotyping: HC JFW. ORCADES Data analysis: CH VV. SHIP Project conception, design and management: SG GH BK HV. SHIP Phenotype collection and data management: SG BK HV. SHIP Genotyping: GH. SHIP Data analysis: SG GH BK HV. TwinsUK Project conception, design and management: TDS GZ. TwinsUK Phenotype collection and data management: MM TDS. TwinsUK Genotyping: N. Soranzo. TwinsUK Data analysis: GZ. Vis Project conception, design and management: HC CH OP IR AFW. Vis Phenotype collection and data management: HC CH OP IR AFW. Vis Genotyping: CH IR AFW. Vis Data analysis: CH VV. BHS Project conception, design and management: LJP. BHS Phenotype collection and data management: GC AWM LJP. BHS Data analysis: GC J. Hui LJP.

## References

- Schunemann HJ, Dorn J, Grant BJ, Winkelstein W, Jr., Trevisan M (2000) Pulmonary function is a long-term predictor of mortality in the general population: 29-year follow-up of the Buffalo Health Study. *Chest* 118: 656–664.
- Rabe KF, Hurd S, Anzueto A, Barnes PJ, Buist SA, et al. (2007) Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *American Journal of Respiratory & Critical Care Medicine* 176: 532–555.
- Hubert HB, Fabsitz RR, Feinleib M, Gwinn C (1982) Genetic and environmental influences on pulmonary function in adult twins. *American Review of Respiratory Disease* 125: 409–415.
- Laurell C-B, Eriksson S (1963) The Electrophoretic alpha 1-Globulin Pattern of Serum in alpha 1-Antitrypsin Deficiency. *Scandinavian Journal of Clinical and Laboratory Investigation* 15: 132–140.
- Hersh CP, Demeo DL, Lange C, Litonjua AA, Reilly JJ, et al. (2005) Attempted replication of reported chronic obstructive pulmonary disease candidate gene



- associations. *American Journal of Respiratory Cell & Molecular Biology* 33: 71–78.
6. Smolonska J, Wijmenga C, Postma DS, Boezen HM (2009) Meta-analyses on suspected chronic obstructive pulmonary disease genes: a summary of 20 years' research. *American Journal of Respiratory & Critical Care Medicine* 180: 618–631.
  7. Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, et al. (2010) Genome-wide association study identifies five loci associated with lung function. *Nature Genetics* 42: 36–44.
  8. Hancock DB, Eijgelsheim M, Wilk JB, Gharib SA, Locher LR, et al. (2010) Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nature Genetics* 42: 45–52.
  9. Wilk JB, Chen T-H, Gottlieb DJ, Walter RE, Nagle MW, et al. (2009) A genome-wide association study of pulmonary function measures in the Framingham Heart Study. *PLoS Genetics* 5: e1000429.
  10. Pillai SG, Ge D, Zhu G, Kong X, Shianna KV, et al. (2009) A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genetics* 5: e1000421.
  11. Kalsheker NA (2009) alpha1-Antitrypsin deficiency: best clinical practice. *Journal of Clinical Pathology* 62: 865–869.
  12. DeMeo DL, Silverman EK (2004) Alpha1-antitrypsin deficiency. 2: genetic aspects of alpha(1)-antitrypsin deficiency: phenotypes and genetic modifiers of emphysema risk. *Thorax* 59: 259–264.
  13. Chappell S, Daly L, Morgan K, Guetta Baranes T, Roca J, et al. (2006) Cryptic haplotypes of SERPINA1 confer susceptibility to chronic obstructive pulmonary disease. *Human Mutation* 27: 103–109.
  14. Gretarsdottir S, Thorleifsson G, Reynisdottir ST, Manolescu A, Jonsdottir S, et al. (2003) The gene encoding phosphodiesterase 4D confers risk of ischemic stroke.[Erratum appears in *Nat Genet.* 2005 May;37(5):555]. *Nature Genetics* 35: 131–138.
  15. Reneland RH, Mah S, Kammerer S, Hoyal CR, Marnellos G, et al. (2005) Association between a variation in the phosphodiesterase 4D gene and bone mineral density. *BMC Medical Genetics* 6: 9.
  16. Mehats C, Jin SLC, Wahlstrom J, Law E, Umetsu DT, et al. (2003) PDE4D plays a critical role in the control of airway smooth muscle contraction. *FASEB Journal* 17: 1831–1841.
  17. Hansen G, Jin S, Umetsu DT, Conti M (2000) Absence of muscarinic cholinergic airway responses in mice deficient in the cyclic nucleotide phosphodiesterase PDE4D. *Proceedings of the National Academy of Sciences of the United States of America* 97: 6751–6756.
  18. Homma S, Sakamoto T, Hegab AE, Saitoh W, Nomura A, et al. (2006) Association of phosphodiesterase 4D gene polymorphisms with chronic obstructive pulmonary disease: relationship to interleukin 13 gene polymorphism. *International Journal of Molecular Medicine* 18: 933–939.
  19. Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, et al. (2009) Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. *American Journal of Human Genetics* 84: 581–593.
  20. Marchini J, Howie B, Myers S, McVean G, Donnelly P (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 39: 906–913.
  21. Devlin B, Roeder K (1999) Genomic control for association studies. *Biometrics* 55: 997–1004.
  22. Li J, Ji L (2005) Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* 95: 221–227.
  23. Palmer LJ, Knuiman MW, Divitini ML, Burton PR, James AL, et al. (2001) Familial aggregation and heritability of adult lung function: results from the Busseton Health Study. *European Respiratory Journal* 17: 696–702.