Journal of Genetics (2019) 98:76
https://doi.org/10.1007/s12041-019-1119-9

REVIEW ARTICLE

Genetic underpinnings of lung function and COPD

ASTHA RANJAN[1], AMANJOT SINGH[2], GAGANDEEP KAUR WALIA[3], MOHINDER PAL SACHDEVA[1] and VIPIN GUPTA[1]*

[1] *Department of Anthropology, University of Delhi, Delhi 110 007, India*
[2] *Department of Anthropology, Panjab University, Chandigarh 160 014, India*
[3] *Public Health Foundation of India, Delhi-NCR, Gurgaon 122 002, India*
*For correspondence. E-mail: udaiig@gmail.com.

Received 8 October 2018; revised 23 April 2019; accepted 10 May 2019

Abstract. Spirometry based measurement of lung function is a global initiative for chronic obstructive lung disease (GOLD) standard to diagnose chronic obstructive pulmonary disease (COPD), one of the leading causes of mortality worldwide. The environmental and behavioural risk factors for COPD includes tobacco smoking, air pollutants and biomass fuel exposure, which can induce one or more abnormal lung function patterns. While smoking remains the primary risk factor, only 15–20% smokers develop COPD, indicating that the genetic factors are also likely to play a role. According to the study of Global Burden of Disease 2015, ~174 million people across the world have COPD. From a comprehensive literature search conducted using the 'PubMed' and 'GWAS Catalogue' databases, and reviewing the literature available, only a limited number of studies were identified which had attempted to investigate the genetics of COPD and lung volumes, implying a huge research gap. With the advent of genomewide association studies several genetic variants linked to lung function and COPD, like *HHIP*, *HTR4*, *ADAM19* and *GSTCD* etc., have been found and validated in different population groups, suggesting their potential role in determining lung volume and risk for COPD. This article aims at reviewing the present knowledge of the genetics of lung function and COPD.

Keywords. genomewide association study; genetics; spirometry; lung function; lung volumes; chronic obstructive pulmonary disease.

Introduction

Lung function, a strong indicator of the physiological condition of the airways and lungs (Wilk *et al.* 2000), is globally considered a significant long-term predictor of morbidity and mortality (Hole *et al.* 1996; Schünemann *et al.* 2000; Myint *et al.* 2005). During the early adulthood, lung volume peaks, then plateaus, and subsequently declines, all these stages are probably affected by both genetic and environmental factors (Young *et al.* 2007). Chronic obstructive pulmonary disease (COPD) is a common disease but is preventable and treatable. It is characterized by the persistent respiratory symptoms and airflow limitations resulting from airway and/or alveolar abnormalities, normally induced by significant exposure to noxious particles or gases (GINA, https://ginasthma.org/#). The global burden of disease (GBD) 2013 study predicts that by 2020, of the 68 million deaths worldwide, 11.9 million will result from lung diseases, while COPD alone will be responsible for 4.7 million deaths (Naghvi *et al.* 2015). According to the GBD 2015 study, globally around 174 million people were estimated to have COPD. They also estimated that shifts in population age structures accounted for 24.2% increase in mortality due to COPD between 2005 and 2015 (Wang *et al.* 2016).

Spirometry is the standard procedure to measure lung function, identify respiratory illness, like COPD, and monitor lung deterioration (Ong *et al.* 2013). The primary estimations of lung volumes are: (i) forced expiratory volume in 1 s (FEV_1) is the volume of air exhaled in the first second and is frequently utilized as an index measure of airway obstruction (Rabe *et al.* 2007). It enables the severity of the disease to be classified and the progression of obstructive lung disease to be followed over time (Rabe *et al.* 2007; Hancock *et al.* 2010). (ii) Forced vital capacity (FVC) is the amount of air which can be forcibly exhaled from the lungs after taking in the deepest breath possible (Boone 2014). It gives the approximate vital

Published online: 31 July 2019

capacity (Rabe *et al.* 2007). A reduced FVC value implies restrictive ventilatory defect, and behaves as an independent (of age and smoking) and reliable mortality predictor in the human population (Hancock *et al.* 2010; Zappala *et al.* 2010; Raghu *et al.* 2011). (iii) FEV_1/FVC: The FEV_1 to FVC ratio is the proportion of the vital capacity of an individual, which can be exhaled in the first second of forced expiration (Rabe *et al.* 2007). The reduced ratio is suggestive of airflow obstruction and used in the diagnosis of various respiratory diseases, regardless of lung size (Strachan 1992; Schünemann *et al.* 2000; Myint *et al.* 2005; Young *et al.* 2007; Hancock *et al.* 2010).

Lung function is regarded as a complex phenotype affected by multiple genetic and environmental factors, and their interactions (Liao *et al.* 2014). The functional genomics research on foetal lung cells has reported that the events in early life play a role in this, as well in future risk for COPD (Hobbs *et al.* 2017). All the lung function genomewide association study (GWAS) signals collectively clarify 9.6% of the heritability for FEV_1, 6.4% of the heritability for FVC and 14.3% of the heritability for FEV_1/FVC in the studied populations (Wain *et al.* 2017).

Demonstration of airflow obstruction with postbronchodilator spirometry is required for a formal diagnosis of COPD, as the patients with partly or fully reversible airflow obstruction may have different pathological processes, contributing to airflow obstruction and can be misclassified as asthma or symptoms of both COPD and asthma may exist in the same individual. But this inclusion of patients with asthma can overestimate the contribution of the SNPs to COPD. However, few studies showed that the findings were not altered substantially even if they excluded asthma patients in their study (Artigas *et al.* 2011a; Hobbs *et al.* 2017). This work includes a review of all the GWAS related to COPD and lung function performed worldwide, along with their validation studies.

Methodology

A comprehensive literature search was conducted using 'PubMed' and 'GWAS Catalogue' databases with the search terms 'GWAS' OR 'genetics' OR 'validation' OR 'gene expression studies' and 'COPD' OR 'Lung function' OR 'FEV_1' OR 'FVC' OR 'FEV_1/FVC', while preparing this review. The literature available in the English language, with a bearing on COPD, lung function phenotypes, GWAS studies, candidate gene studies, genetic association studies, validation studies, gene expression studies, and gene–environment studies were included in this review. A total of 72 studies which satisfied the criteria were selected. The genetic regions revealing relevance at the genomewide significance level ($P < 5 \times 10^{-8}$) were then identified for this review.

GWAS

GWAS screenes thousands of single-nucleotide polymorphisms (SNPs) to identify the potential genetic variants related to a given phenotype (Gibson 2010). Until now, 17 GWASs of COPD and lung function measures have been conducted in the Western populations, some of these studies have even conducted meta-analyses along with their study, and 122 loci significantly associated with COPD and lung function were identified, where most of these reviewed GWASs reported novel findings. Among these, 15 were common variants, namely *HHIP*, *HTR4*, *ADAM19*, *FAM13A*, *RARB*, *GSTCD*, *TGFβ2*, *CHRNA3/5*, *CDC123*, *MMP3/12*, *AGER*, *RIN3*, *ARMC2*, *CYP2A6* and *EEFSEC* (tables 1&2, 4–6). All these studies have been performed using both the gender, ≥ 40 years as an average age on ever/never smokers, except for two studies, Cho *et al.* (2010) and Artigas *et al.* (2011b), all the others have performed imputation of SNPs. The genomic inflation factor ranged from 0.946 (Wilk *et al.* 2012) to 1.372 (Wain *et al.* 2017) for all the reviewed GWASs. It has been suggested that numerous undetected common variants are present, those of modest effect sizes and of low frequency (1% <MAF≤5%) and those which are rare (MAF≤1%) and of greater effect sizes, accounting for the 'missing heritability' (Manolio *et al.* 2009; Gibson 2011). It is essential to validate the GWAS loci in different populations, to identify the universal biomarkers and potential future intervention strategies, particularly in the countries supporting diverse populations and carrying a heavy burden of the disease.

Artigas *et al.* (2015) conducted a GWAS on 38,199 European ancestry individuals, followed-up to top associations on 54,550 individuals and discovered 16 new lung function signals. Despite having >80% power in discovery to detect associations for variants with MAF of 0.5 and 1%, and effect sizes above 0.3 and 0.2 s.d. units, respectively, none of the associations with lower allele frequency variants had been detected in the study. This study shows that the study of lung function measures is a powerful approach to bring insights into the genetics of COPD (Artigas *et al.* 2015). To date, Hobbs *et al.* (2017) have conducted the largest GWAS of COPD cases with over 60,000 subjects (including 15,256 COPD cases) and combined subjects of different ethnicities, hypothesizing that shared COPD risk factors across ethnicities would outweigh power loss due to heterogeneity. The majority of their significant loci overlapped with the lung function loci, strengthening the foundation for investigating the relationship of lung function variability in the general population to risk of developing COPD (Hobbs *et al.* 2017). Wain *et al.* (2017) provided the most comprehensive evidence regarding genetic variants associated with lung function and their association with susceptibility to COPD, with a more than threefold difference in

Table 1. Genetic variants identified by GWAS related to COPD.

Population	Sample size	Gene	Ch. no.	SNP	OR (95% CI)	*P* value	Reference
European	Cases: 823 Controls: 810	*ANKH*	5	rs7727670	1.86 (1.48–2.34)	8.38×10^{-8}	Pillai *et al.* (2009)
				rs9686327	1.84 (1.47–2.31)	9.84×10^{-8}	
				rs7341022	1.84 (1.47–2.31)	9.97×10^{-8}	
European & American	Cases: 2940 Controls: 1380	*FAM13A*	4	rs1903003	0.73 (0.64–0.81)	7.74×10^{-8}	Cho *et al.* (2010)
European & American	Cases: 3499 Controls: 1922	*FAM13A*	4	rs7671167	0.73 (0.66–0.81)	2.22×10^{-9}	Cho *et al.* (2012)
				rs1964516	0.73 (0.66–0.81)	1.88×10^{-9}	
		–	19	rs7973	0.74 (0.63–0.83)	2.88×10^{-9}	
Non-Hispanic White, African-American, & European	Cases: 6633 Controls: 5704	*TGFB2*	1	rs4846480	1.26 (1.16–1.37)	1.25×10^{-9}	Cho *et al.* (2014)
		FAM13A	4	rs13141641	1.39 (1.36–1.51)	3.66×10^{-15}	
		HHIP	4	rs4416442	1.3 6 (1.26–1.47)	9.44×10^{-15}	
		MMP3/12	11	rs626750	1.36 (1.23–1.51)	5.53×10^{-9}	
		RIN3	14	rs754388	1.33 (1.20 –1.48)	6.69×10^{-8}	
		CHRNA3	15	rs12914385	1.39 (1.28–1.51)	2.70×10^{-16}	
Hispanics	Cases: 447	*KLHL7 & NUPL2*	7	rs858249	*β = −0.84	6.1×10^{-8}	Chen *et al.* (2015)
	Controls:1085	*DLG2*	11	rs286499	*β = −1.2	8.4×10^{-8}	
European & American	Cases:15256 Controls:47936	*TGFB2*	1	rs10429950	1.12 (1.07–1.16)	1.66×10^{-10}	Hobbs *et al.* (2017)
		PID1	2	rs16825267	1.24 (1.15–1.34)	1.68×10^{-9}	
		EEFSEC	3	rs2955083	1.20 (1.12–1.27)	4.16×10^{-13}	
		FAM13A	4	rs6837671	1.16 (1.11–1.20)	7.48×10^{-15}	
		GSTCD	4	rs11727735	1.27 (1.17–1.37)	3.84×10^{-14}	
		HHIP	4	rs13141641	1.23 (1.18–1.28)	9.10×10^{-41}	
		TET2	4	rs2047409	1.10 (1.06–1.15)	2.46×10^{-13}	
		ADAM19	5	rs113897301	1.20 (1.13–1.26)	1.58×10^{-13}	
		HTR4	5	rs7733088	1.18 (1.13–1.23)	5.33×10^{-26}	
		ADGRG6	6	rs9399401	1.14 (1.09–1.19)	1.81×10^{-19}	
		AGER	6	rs2070600	1.28 (1.15–1.41)	5.94×10^{-10}	
		ARMC2	6	rs2806356	1.12 (1.07–1.18)	8.34×10^{-10}	
		DSP	6	rs2076295	1.11 (1.07–1.18)	3.97×10^{-9}	
		RIN3	14	rs754388	1.20 (1.14–1.26)	4.96×10^{-14}	
		CHRNA5	15	rs17486278	1.22 (1.18–1.27)	1.77×10^{-28}	
		THSD4	15	rs14411358	1.13 (1.09–1.18)	8.22×10^{-16}	
		CCDC101	16	rs17707300	1.12 (1.08–1.17)	6.75×10^{-10}	
		CFDP1	16	rs7186831	1.12 (1.07–1.18)	1.12×10^{-11}	
		MTCL1	18	rs647097	1.11 (1.06–1.15)	6.14×10^{-9}	
Hispanics & Latinos	Cases: 363	*PDZD2*	5	rs7709630	−0.607 (0.101)#	1.56×10^{-8}	Burkart *et al.* (2018)
	N = 5616	*CDRT15P1*	11	rs2286351	−0.804 (0.136)#	1.97×10^{-8}	

Ch. no., chromosome number; OR, odds ratio at 95% confidence interval; *β, beta estimate; #standard error.

COPD risk between highest and lowest allelic risk score deciles, but they also reported that the lung function-associated variants were not associated with acute exacerbations of COPD (Wain *et al.* 2017). The studies with largest effect sizes for *GEMIN5 KIF4B* and *FLJ20184* were associated with FEV_1 among the Europeans and Hispanics/Latinos populations (Hancock *et al.* 2010; Burkart *et al.* 2018).

All the reviewed literature on GWASs, validation studies, candidate gene studies, genetic association studies and their findings are listed below and tabulated as per the studied phenotypes: COPD and lung function measures.

COPD

Based on the case–control design, GWASs of COPD have been conducted on the European, European American, Hispanic and African-American populations (table 1). The effect sizes of the identified genetic markers vary from OR = 1.09 to 1.89 among the European and American populations (table 1). The genes identified through these GWAS have been validated in different populations (table 2) and are involved in biological pathways of inflammation, morphogenesis and development (table 3). The population attributable risk for COPD in GWAS

Table 2. Candidate gene association and validation of GWAS loci related to COPD.

Population	Sample size	Gene	Ch. no.	SNP	OR (95%CI)	*P* value	Reference
Korean	Cases: 83 Controls: 203	*SLC11A1*	2	rs17235409 rs1059823	2.23 (1.24–4.02) 1.92 (1.10–3.35)	0.007 0.022	Kim *et al.* (2008)
Chinese	Cases: 34 Controls: 71	*PTEN*	10	rs701848 rs1903858	0.12 (0.03–0.47) 0.13 (0.04–0.44)	0.002 0.0010	Hosgood *et al.* (2009)
Chinese	Cases: 234 Controls: 312	*CD40-1C/T*	20	–	1.77 (1.12–2.83)	0.015	Liu *et al.* (2009)
European (Scottish & German)	Cases: 394 Controls: 3676	*HHIP*	4	rs12504628	1.19 (1.12–1.27)	4.55×10^{-8}	Artigas *et al.* (2011b)
Polish	Cases: 315 Controls: 330	*HHIP*	4	rs13118928	0.68 (0.53 –0.87)	0.002	Zhou *et al.* (2012)
Polish	Cases: 315	*ADCY2*	5	rs11134242	1.35 (1.10–1.70)	0.01	Hardin *et al.* (2012)
	Controls: 330	*CHRNA3/5*	15	rs8034191	1.89 (1.50–2.40)	7.40×10^{-7}	
		IREB2	15	rs13180	0.69 (0.50–0.90)	3.40×10^{-3}	
Han Chinese	Cases: 279 Controls: 367	*ADIPOQ*	3	rs1501299	1.54 (1.11–2.75)	0.009	Yuan *et al.* (2012)
Han Chinese	Cases: 331 Controls: 213	*HHIP* *TNF-α*	4 6	rs2353397 rs1800629	2.16 (1.66–2.81) 1.97 (1.21–3.21)	< 0.0001 0.006	Guo *et al.* (2012)
Han Chinese	Cases: 680 Controls: 687	*HHIP*	4	rs10519717	1.53 (1.08–2.14)	0.0300	Wang *et al.* (2013a)
Han Chinese	Cases: 680 Controls: 687	*FAM13A*	4	rs7671167	0.58 (0.34–0.98)	0.026	Wang *et al.* (2013b)
Chinese Li	Cases: 234 Controls: 240	*FAM13A* *SETD*	4 7	rs7671167 rs17050782	1.58 (1.05–2.38) 2.30 (1.22–4.31)	0.028 0.008	Ding *et al.* (2015)
Mexican (Mestizo)	Cases: 330 Controls: 808	*MMP-2*	16	rs11646643 rs243864	1.58 (1.07–2.34) 7.44 (3.62–5.26)	2.58×10^{-2} 1.0×10^{-10}	Hernández-Montoya *et al.* (2015)
		MMP-9	20	rs3918253	1.72 (1.08–2.71)	2.58×10^{-2}	
Polish	Cases: 617 Controls: 524	*FAM13A*	4	rs2869967	2.41 (1.44–4.05)	0.0007	Ziółkowska-Suchanek *et al.* (2015)

OR, odds ratio at 95% confidence interval; ch. no, chromosome number.

Table 3. Biological pathways involved in COPD and lung function.

Biological pathway	Genes involved	References
Inflammation pathway	*IL16*, *IL6*, *SLC11A1*, *CD40*, *PLUNC*, *HTR4*, *AGER*, *THSD4*, *CHRNA3*/*CHRNA5*/*CHRNB4*, *HDAC AAT*, *MAGI2* and *NT5C3B*	Kim *et al.* (2008), Liu *et al.* (2009), Artigas *et al.* (2011a), Hardin *et al.* (2012), Liao *et al.* (2013), Tang *et al.* (2014), Yuan *et al.* (2012), Wilk *et al.* (2007), Lee *et al.* (2011), Siedlinski *et al.* (2011), Thun *et al.* (2013), Dijkstra *et al.* (2015)
Xenobiotic signalling	*GST* family, *ARNT*, *IL6*, *ALDH8A1*, *TRIP11*, *PPARGC1*, *PRKCE*	Wilk *et al.* (2009), Hancock *et al.* (2010), Zhou *et al.* (2012)
Morphogenesis or development pathway (lung development)	*NPNT*, *NCR3*/*AIF1*, *HHIP*, *FOXA1*, *TCF21*, *SOX9* and *MYBPC1*	Artigas *et al.* (2011a), Hancock *et al.* (2010, 2012), Pillai *et al.* (2009), Hansel *et al.* (2013), Brandsma *et al.* (2015)
Lung tissue repair pathway	*PPAP2B*, *BMP6*, *FBLN5*, *LTBP2*, *MFAP4*, *TGFB2*, *MMP15* and *EFEMP1*	Loth *et al.* (2014), Wilk *et al.* (2012)
Smooth muscle contraction and airway constriction pathway	*MYH11*, *SERPIN* family, *TMEM26*, *PPARGC1A*, *MYH15* and *SGCD*	Wilk *et al.* (2009), Wang *et al.* (2013a, b), Siedlinski *et al.* (2011), Hansel *et al.* (2013)

ranges from ~ 5% (Cho *et al.* 2012) to 48% (Wain *et al.* 2017).

FEV$_1$

GWASs of FEV_1 have been conducted in the European population using the population-based cohort research design. The effect sizes of the genetic variants identified range from –57.90 mL to 70.08 mL per allele change for FEV_1 (table 4). The gene regions associated with FEV_1 play a role in biological pathways of inflammation, morphogenesis, development and xenobiotic signalling (table 3). The loci thus identified account for the 0.07% (Rabe *et al.* 2007) to 4% (Artigas *et al.* 2015) genetic variations in the FEV_1 among the European population.

A candidate gene study performed in the Polish population (n = 728 and controls = 452) reported the association of *IL8* and a decline in the FEV_1 over time in COPD patients: rs4073 (OR = 1.43 (1.05–1.88, P = 0.047)); rs2227306 (OR = 1.46 (1.06–1.94, P = 0.041)); and rs2227307 (OR = 1.46 (1.08–1.87, P = 0.011)) (Córdoba-Lanús *et al.* 2015).

A very recent GWAS on postbronchodilator lung function measures among 1086 non-Hispanic White European descent smokers showed significant association of *SERPINA1* gene variant rs28929474 ($\beta = -13.6$, $P = 3.5 \times 10^{-8}$) with postbronchodilator per cent predicted FEV_1(SE not given). This study is the first to show association of *SERPINA1* gene with lung function, it has been previously associated with COPD phenotype (Li *et al.* 2018).

FVC

GWASs of FVC have been conducted among the European population using the population-based cohort research design. The effect sizes are in the range of –21.12 mL to 30.88 mL per allele change for the FVC (table 5). The gene regions related to the FVC are involved in the lung tissue repair pathway (table 3). The genetic loci identified explain the 0.74% (Loth *et al.* 2014) to 3.20% (Artigas *et al.* 2015) of the variations among the European population.

A candidate gene study on the variants of the *IL8* gene on chromosome 4 has been performed in the Polish population and the following associations of rs4073 (OR = 2.77 (2.29–3.54, P = 0.008)); rs2227306 (OR = 2.76 (2.21–3.54, P = 0.051)) and rs2227307 ((OR = 2.78 (2.28–3.56, P = 0.015)) were noted, with a decline in the FVC over time (Córdoba-Lanús *et al.* 2015).

FEV$_1$/FVC ratio

GWASs of FEV_1/FVC have been performed among the European, African-American and Hutterite populations using the population-based cohorts. The effect size ranged from 1.0 to 1.10 among the Europeans (table 6). The gene regions affecting the FEV_1/FVC ratio are involved in the inflammation and morphogenesis or development pathways (table 3). The loci thus identified were responsible for 0.07% (Repapi *et al.* 2010) to 5.4% (Artigas *et al.* 2015) variance among the populations studied, as demonstrated by these genetic regions. The *AP1S2* locus was the first to be recorded on the X-chromosome for lung function (Artigas *et al.* 2015). Recently, two low-allele frequency variants *LTBP4* (MAF = 1.5%) and *GPR126* (MAF = 2.4%), respectively, were detected in association with FEV_1/FVC (Artigas *et al.* 2015).

In a validation study performed among the Polish population involving 645 individuals (Caucasian controls = 330), two loci were found to be associated with FEV_1/FVC i.e. rs7671167 in *FAM13A* (β = 1.39 (0.41), P = 0.0008) and rs11134242 in ADCY2 ($\beta = -2.50$ (1.20), P = 0.03) (Han *et al.* 2010).

Li *et al.* (2018) also showed significant association of *SERPINA1* gene variant rs28929474 ($\beta = -0.087$, $P = 1.2 \times 10^{-8}$) with postbronchodilator FEV_1/FVC (SE not given).

Overlapping signals

The GWAS loci found in association with lung function and COPD are also recognized as linked to other traits and a variety of chronic disorders (Wain *et al.* 2017), implying either the pleiotropic effects of these loci or common underlying biological pathways. The variants of *DSP*, *FAM13A* (posterior probability of association >0.99) and *MAPT-KANSL1* (posterior probability of association >0.84) have been indicated to be overlapping signals with pulmonary fibrosis and COPD, and the variants associated with pulmonary fibrosis were observed to exert a concordant direction of effect in COPD (Hobbs *et al.* 2017). A total of 17 loci associated with COPD are found to overlap with the loci for lung function, strengthening the grounds for the investigation of the relationship between the variability in lung function in the general population and risk of developing COPD (Hardin *et al.* 2012; Hobbs *et al.* 2017).

Further, a few of the loci associated with lung function decline overlapped with other traits and diseases. The genes related to FEV_1/FVC such as *KCNE2* with myocardial infarction revealed identical direction of effect (Artigas *et al.* 2011a). Similarly, genes associated with the FEV1 like *RIN3* are also reported to be linked to Paget's disease and bone mineral density, while *NCR3* and *ZKSCAN3*/*ZNF323* were related to lung cancer in the same direction of effect (Artigas *et al.* 2011a).

Table 4. Genetic variants identified by GWAS related to FEV_1.

Population	Sample size	Gene	Ch. no.	SNP	β(S.E.)	*P* value	Reference
European	74,564	*TNS1*	2	rs2571445	0.03 (0.005)	1.11×10^{-12}	Pillai *et al.* (2009)
		GSTCD	4	rs10516526	0.08(0.009)	2.18×10^{-23}	
		HHIP	4	rs12504628	−0.06 (0.01)	1.50×10^{-10}	
		HTR4	5	rs6889822	0.03(0.006)	8.17×10^{-9}	
				rs3995090	0.038(0.006)	4.29×10^{-9}	
European	20,890	*FLJ20184*	4	rs17036052	70.08	1.83×10^{-15}	Hancock *et al.* (2010)*
				rs17035960	53.85	9.42×10^{-14}	
		GSTCD	4	rs11097901	59.14	3.26×10^{-16}	
				rs11728716	57.47	7.20×10^{-16}	
		HHIP	4	rs1980057	21.04	5.86×10^{-9}	
		INTS12	4	rs11727189	64.70	4.66×10^{-17}	
				rs17036090	−57.90	5.61×10^{-15}	
		NPNT	4	rs17331332	56.79	5.69×10^{-15}	
				rs17036341	−56.65	2.18×10^{-15}	
		HTR4	5	rs11168048	−21.49	6.69×10^{-9}	
European	67,171–92,977	*ZKSCAN3/ZNF323*	6	rs6903823	−0.03(0.006)	2.18×10^{-10}	Artigas *et al.* (2011a)
		CDC123	10	rs1878798	0.02(0.004)	1.84×10^{-9}	
				rs7068966	0.02(0.004)	2.82×10^{-12}	
		C10orf11	10	rs11001819	−0.02 (0.004)	2.98×10^{-12}	
European	38,199	*MCL1*	1	rs6681426	0.02(0.005)	4.35×10^{-9}	Artigas *et al.* (2015)
		TBX3	12	rs10850377	−0.03 (0.005)	2.50×10^{-12}	
		TRIP11	14	rs7155279	−0.03 (0.005)	1.41×10^{-9}	
		MIAT	22	rs134041	−0.02 (0.005)	3.03×10^{-9}	
European (non-Hispanic Whites) & African Americans	13,532	*LOC646576*	4	rs138641402	−0.08	8.99×10^{-15}	Wilk *et al.* (2012)*
		(near *HHIP*)					
		FAM13A	4	rs6837671	0.06	2.89×10^{-13}	
		DBH	9	rs1108581	−0.05	8.72×10^{-9}	
		CHRNA3	15	rs56077333	−0.08	5.29×10^{-18}	
		AGPHD1	15	rs8031948	−0.07	2.86×10^{-15}	
		CHRNA5	15	rs17486278	0.07	3.48×10^{-18}	
		CHRNβ4	15	rs17487223	−0.07	2.06×10^{-14}	
		IREB2	15	rs2568494	−0.06	1.66×10^{-12}	
		PSMA4	15	rs58365910	0.07	6.69×10^{-15}	

Table 4 (*contd*)

Population	Sample size	Gene	Ch. no.	SNP	β (S.E.)	*P* value	Reference
European	48,943	*ABLIM3*	5	rs38392434	−0.02 (0.004)	4.48×10^{-11}	Wain *et al.* (2017)
		HLA-DQB1/HLA-DQA2	6	rs114229351	−0.03 (0.006)	2.12×10^{-10}	
		LOC389602-LOC285889	7	rs12698403	−0.02(0.004)	1.11×10^{-13}	
		GLIS3	9	rs7872188	−0.02 (0.004)	1.59×10^{-10}	
		AHNAK	11	rs2509961	0.03 (0.004)	1.49×10^{-13}	
		CDON-RPUSD4	11	rs567508	0.03 (0.005)	4.77×10^{-10}	
		ME3-PRSS23	11	rs145729347	−0.03 (0.005)	8.58×10^{-9}	
		MSRB3	12	rs1494502	0.02(0.004)	9.80×10^{-10}	
		ZGPAT	20	rs72448466	−0.03 (0.004)	4.31×10^{-12}	
		MICAL3	22	rs11704827	0.03(0.005)	8.32×10^{-13}	
		MN1	22	rs2283847	−0.02 (0.004)	3.41×10^{-11}	
Hispanics Latinos	11,822	*ZSWIM7* (intron)	17	rs4791658	33.35(5.70)	4.99×10^{-9}	Burkart *et al.* (2018)
		GEMIN5 KIF4B	5	rs115745680	313.24(57.32)	4.63×10^{-8}	

β, Beta estimate, mL change per allele; S.E., standard error; ch. no., chromosome number; *standard error of effect size not mentioned.

Biological pathways

A thorough review of the GWASs has revealed that around 90 genes in association with crucial biological pathways were related to COPD and lung function; however more than 40 genes that are identified to be associated with COPD and lung function measures through GWASs, as yet, have not been studied in terms of their biological functions and roles. Overall, the loci identified in the GWASs are involved in xenobiotic handling, metabolism signalling, inflammation, growth and development, etc., and contribute towards the development of COPD (Wilk *et al.* 2009; Repapi *et al.* 2010; Tang *et al.* 2014). Interestingly, some of the biological pathways are directly related to lung development, tissue repair and inflammation, as well as smooth muscle contraction and airway constriction (see table 3).

Gene–environment interaction studies

Some study has been conducted to observe the results of the interaction between specific environmental factors and the genetic effects on COPD and lung function. Genotype-by-environment interaction between the risk of the rs8034191 genotype and current smoking status on COPD has been observed in the Norwegian sample ($P = 0.002$), which shows a substantially higher risk of COPD in current smokers carrying the rs8034191 C allele (OR = 2.0) than in former smokers (OR = 1.1) (Pillai *et al.* 2009). The C allele of rs360563 at *CRISP2* has been confirmed to accelerate the decline in FEV_1/FVC by 1.1% per interquartile range (IQR) change in the particulate matter with diameter of 10 μm (PM_{10}) exposure over 11 years (Curjuric *et al.* 2012). Similarly, the G allele of rs2035268 at *SNCA* was also shown to accelerate the decline in the FEV_1/FVC by 3.8% per allele and the IQR change in PM_{10} exposure (Curjuric *et al.* 2012). Further, the effect of the interaction of the *SLC38A8* at rs9931086 with occupational exposure on the FEV_1 suggests the mediating pathways caused by the gene–environment interactions (Liao *et al.* 2013).

Gene expression studies

Whole-genome gene expression studies of the diseased lung and airway tissues facilitate a clear understanding of the processes that contribute to disease pathogenesis that can pave the way for the development of novel therapeutic approaches (Silverman *et al.* 2009). The literature review reveals that only a very few gene expression studies were conducted on COPD and lung-function-related genes, to date. While remarkable heterogeneity is evident among the studies, most of the lung tissue and peripheral blood-related research studies have identified the enrichment of differentially expressed genes in the inflammatory pathways related to immune regulation, particularly the

Table 5. Genetic variants identified by GWAS related to FVC.

Population	Sample size	Gene	Ch. no.	SNP	β (S.E.)	*P* value	Reference
European	85,170	*EFEMP1*	2	rs1430193	−21.12 (2.99)	1.86×10^{-12}	Loth *et al.* (2014)
		BMP6	6	rs6923462	30.88 (4.28)	5.89×10^{-13}	
		PRDM11	16	rs2863171	23.92 (3.90)	8.97×10^{-10}	
		WWOX	16	rs1079572	16.25 (2.83)	9.95×10^{-9}	
		KCNJ2	17	rs6501431	23.05 (3.88)	2.94×10^{-9}	
European	38,199	*AK097794*	3	rs6441207	0.03 (0.005)	1.27×10^{-13}	Artigas *et al.* (2015)
		LHX3	9	rs2274116	0.03 (0.005)	5.55×10^{-14}	
		PTHLH	12	rs11383346	−0.04 (0.005)	9.52×10^{-18}	
European	48,943	*SPAG17- TBX15*	1	rs2001534334	0.03 (0.005)	8.20×10^{-14}	Wain *et al.* (2017)
		SUCLG2	3	rs1490265	0.02 (0.004)	1.58×10^{-9}	
		LOC340113-TARS	5	rs91731	−0.04 (0.007)	4.31×10^{-13}	
		ARL15	5	rs2441026	0.02 (0.004)	9.32×10^{-10}	
		GLIS3	9	rs7872188	−0.02 (0.004)	9.32×10^{-10}	
		MYPN	10	rs70955607	−0.03 (0.004)	8.67×10^{-15}	
		TBX3- MED13L	12	rs35506	0.02 (0.004)	9.87×10^{-10}	
		CASC20- BMP2	20	rs6140050	0.03 (0.004)	6.39×10^{-14}	

β, Beta estimate, mL change per allele; S.E., standard error; Ch. no., chromosome number.

B-cell and T-cell development and differentiation (Chang *et al.* 2016).

A study performed among Japanese participants showed that the TGF-β1 expression was higher in both smokers and patients with COPD when compared with the nonsmokers (Takizawa *et al.* 2001). Another study demonstrated that the *TLR4* and *HBD2* expressions were increased in patients with mild-to-moderate COPD when compared with the healthy controls (MacRedmond *et al.* 2007). Further, the *FOXO3A* gene expression decreased as airway obstruction progressed, whereas the *CXADR* and *COX2* gene expressions increased in the lungs of the obstructed patients and their degree of expression was observed to rise in proportion to the severity of the disease (Wang *et al.* 2008). The *CD31* and *CD34* gene expressions were noted to decrease in patients with moderate COPD, but to escalate in cases of severe COPD, when compared with the patients who smoked but did not develop COPD (Kato *et al.* 2016). The expression levels of *IL-17*, *IL-10*, *IL-6* and *TGFβ* were observed to increase among the COPD patients exposed to mustard gas (Farahani *et al.* 2017). In India, an increase in *TLR-2* and *TLR-4* gene expressions was recorded among the patients with moderate and severe COPD when compared with the controls (Tripathi *et al.* 2017). It has been reported that the *IL-1* pathway signalling genes, i.e. *IRAK2*, *IRAK3* and *PELI1*, as well as the signalling receptor *IL1R1* showed strong expressions in patients with COPD and asthma (Baines *et al.* 2017). A very recent study reported that the risk allele (T) of COPD SNP variant rs7973 was associated with the increased expression of *EGLN2* gene in blood among the discovery sample population (Nedeljkovic *et al.* 2018).

Limitations and future directions

Unexplained heritability is now a highly recognized challenge in genetic epidemiology (Eichler *et al.* 2010) and possible explanations for it may include, pleiotropy, i.e. multiple effects of the common variants, the roles played by the undetected rare variants, possible gene–environment interactions, gene–gene interactions and epigenetic regulation-mechanisms that are not captured by the existing GWAS platforms (Loth *et al.* 2014). Although, the GWAS revolution commenced a decade ago, most of the GWAS on lung function have been performed on European populations. The fundamental basis for conducting the GWAS was to detect the common genetic variants. The low-frequency or rare-frequency variants often remained undetected in most studies due to poor imputation quality (Artigas *et al.* 2015). While a few of the associated polymorphisms are potentially functional, most of them occur in the intronic regions and are a possible tag for the functional variants that are yet to be identified (Hancock *et al.* 2010). Gene–gene interactions or gene–environment interaction studies on respiratory health are scarce because they necessitate large sample size, robust measurement of exposure and validated genetic variants (Liao *et al.* 2013). There are a limited number of GWASs on COPD and lung function when compared with other diseases such as asthma, lung cancer, etc. Among the currently reviewed 17 GWAS studies, 11 have performed bronchodilator reversibility (prebronchodilator and postbronchodilator response) on their study, which is an important parameter to check for the lung function reversibility in a population (Pillai *et al.* 2009; Cho *et al.* 2010, 2012, 2014; Hansel *et al.* 2013; Chen

Table 6. Genetic variants identified by GWAS related to FEV_1/FVC.

Population	Sample size	Gene	Ch. no.	SNP	β(S.E.)	*P* value	Reference
European	74,564	*HHIP*	4	rs12504628	−0.07(0.01)	6.48×10^{-13}	Pillai *et al.* (2009)
		AGER	6	rs2070600	0.08(0.01)	3.07×10^{-15}	
		DAAM2	6	rs2395730	0.04(0.008)	7.98×10^{-8}	
		THSD4	15	rs12899618	0.06(0.008)	7.24×10^{-15}	
European	20,890	*HHIP*	4	rs1980057	0.52	3.21×10^{-20}	Hancock *et al.* (2010)*
				rs1032295	−0.47	4.37×10^{-15}	
		ADAM19	5	rs2277027	0.38	9.93×10^{-11}	
				rs1422795	0.37	2.62×10^{-10}	
		HTR4	5	rs11168048	−0.40	1.08×10^{-11}	
				rs7735184	0.37	6.23×10^{-11}	
		AGER	6	rs2070600	1.00	3.15×10^{-14}	
		GPR126	6	rs3817928	−0.42	1.17×10^{-9}	
				rs7776375	−0.37	6.71×10^{-9}	
		PPT2	6	rs10947233	1.10	6.66×10^{-12}	
European	63,564–93,997	*MFAP*	1	rs2284746	−0.04(0.005)	7.50×10^{-16}	Artigas *et al.* (2011a)
		HDAC4	2	rs12477314	0.04(0.006)	1.68×10^{-12}	
		RARB	3	rs1529672	−0.04 (0.006)	3.97×10^{-14}	
		NCR3	6	rs2857595	0.03(0.006)	2.28×10^{-10}	
		ARMC2	6	rs2798641	−0.04 (0.007)	8.35×10^{-9}	
		CDC123	10	rs7068966	0.03(0.006)	6.13×10^{-13}	
		–	10	rs1878798	0.03(0.005)	9.56×10^{-11}	
		CCDC38	12	rs1036429	0.03(0.006)	2.30×10^{-11}	
		CFDP1	16	rs2865531	0.03(0.005)	1.77×10^{-11}	
European	50,047	*DNER*	2	rs7594321	−0.01(0.01)	2.64×10^{-9}	Hancock *et al.* (2012)
		HLA-DQB1/HLA-DQA2	6	rs7764819	−0.001(0.02)	4.39×10^{-9}	
European American	4,048	–	10	rs10761570	−9.80	6.0×10^{-8}	Hansel *et al.* (2013)*
		–	14	rs177852	−9.30	6.30×10^{-8}	
European	38,199	*LYPLAL1*	1	rs201204531	−0.03(0.005)	2.68×10^{-10}	Artigas *et al.* (2015)
		KCNS3	2	rs61067109	−0.04 (0.006)	1.40×10^{-15}	
		NPNT	4	rs6856422	−0.05 (0.005)	1.51×10^{-23}	
		GPR126	6	rs148274477	−0.16 (0.015)	9.58×10^{-26}	
		ASTN2	9	rs34886460	0.03(0.005)	4.72×10^{-11}	
		EMP2	16	rs12149828	0.04(0.007)	7.65×10^{-10}	
		LTBP4	19	rs113473882	−0.15 (0.021)	9.95×10^{-13}	
European (non-Hispanic Whites) & African Americans	13,532	*TGFβ2*	1	rs72738834	−0.01	6.51×10^{-9}	Lutz *et al.* (2015)*
		FAM13A	4	rs6837671	0.01	5.45×10^{-13}	

Table 6 (*contd*)

Population	Sample size	Gene	Ch. no.	SNP	β(S.E.)	*P* value	Reference
		LOC646576 (near *HHIP*)	4	rs13141641	−0.01	9.52×10^{-20}	
		MMP12	11	rs72981684	0.01	3.92×10^{-10}	
		MMP3	11	rs72981675	0.01	3.65×10^{-10}	
		RIN3	14	rs754388	−0.01	5.54×10^{-9}	
		CHRNA3	15	rs56077333	−0.01	9.55×10^{-20}	
		AGPHD1	15	rs8042849	0.01	2.81×10^{-15}	
		CHRNA5	15	rs17486278	0.01	4.48×10^{-20}	
		PSMA4	15	rs58365910	0.01	1.24×10^{-15}	
		CHRNβ4	15	rs17487223	−0.01	3.28×10^{-15}	
		IREB2	15	rs17484524	0.01	4.83×10^{-13}	
		CYP2A6	19	rs56113850	0.01	5.19×10^{-9}	
European	48,943	*LOC101929516*	1	rs17513135	−0.03(0.005)	2.31×10^{-16}	Wain *et al.* (2017)
		CDC7-TGFβR3	1	rs1192404	−0.04 (0.005)	6.09×20^{-20}	
		TGFβR3-BRDT2	1	rs12140637	−0.02 (0.004)	1.18×10^{-9}	
		CHRM3	1	rs6688537	−0.03 (0.004)	6.72×10^{-22}	
		TRAF3IP1-ASB1	2	rs61332075	0.03(0.006)	2.55×10^{-10}	
		CACNA2D3-WNT5A	3	rs1458979	−0.02 (0.004)	4.42×10^{-10}	
		EEFSEC	3	rs2811415	−0.03 (0.005)	5.52×10^{-11}	
		LOC100507661- MECOM	3	rs56341938	0.02(0.004)	4.52×10^{-14}	
		FAM13A	4	rs13110699	−0.03 (0.005)	7.86×10^{-15}	
		ITAG1	5	rs1551943	−0.04 (0.007)	4.31×10^{-13}	
		C5orf56	5	rs7713065	0.02(0.004)	2.77×10^{-11}	
		CYFIP2	5	rs10515750	−0.05 (0.007)	5.26×10^{-13}	
		LST1	6	rs28986170	0.06(0.010)	1.56×10^{-10}	
		KCNQ5	6	rs141651520	0.04(0.005)	9.93×10^{-18}	
		ZKSCAN1	7	rs72615157	0.03(0.005)	1.98×10^{-9}	
		SVIL-KIAA1462	10	rs3847402	−0.02 (0.004)	7.72×10^{-11}	
		FGD6	12	rs113745635	−0.04 (0.005)	8.46×10^{-16}	
		MGA	15	rs72724130	−0.05 (0.009)	9.58×10^{-10}	
		THSD4	15	rs12591467	0.02 (0.004)	5.65×10^{-10}	
		SH3GL3	15	rs66650179	−0.04 (0.006)	3.71×10^{-12}	
		EFCAB5	17	rs62070270	−0.03 (0.004)	7.29×10^{-18}	
		CISD3	17	rs11658500	−0.03 (0.006)	7.22×10^{-11}	
Hispanics Latinos	11,822	*KIF25* (intron)	6	rs 76656601	4.70(0.77)	1.31×10^{-9}	Burkart *et al.* (2018)
		HAL (intron)	12	rs145174011	4.85 (0.85)	9.59×10^{-9}	
		KCNE2	21	rs28593428	0.80 (0.14)	1.45×10^{-8}	
		GPR126	6	rs262113	20.57(0.10)	2.83×10^{-8}	
		LOC105375250 (intron)	7	rs74444778	6.17 (1.12)	3.61×10^{-8}	

β, Beta estimate, mL change per allele; S.E., standard error; Ch. no., chromosome number; *standard error of effect size not mentioned.

et al. 2015; Lutz *et al.* 2015; Hobbs *et al.* 2017; Wain *et al.* 2017; Burkart *et al.* 2018; Li *et al.* 2018) whereas, six have performed analysis using prebronchodilator data only (Hancock *et al.* 2010, 2012; Artigas *et al.* 2011a, b, 2015; Wilk *et al.* 2012; Loth *et al.* 2014). Three studies have not reported whether the investigated polymorphisms were in the Hardy–Weinberg equilibrium or not, which is an important quality control measure (Pillai *et al.* 2009; Artigas *et al.* 2011b; Wilk *et al.* 2012). Chen *et al.* (2015) has used beta estimate (which is used for quantitative traits) for effect size estimate of COPD instead of odds ratio (OR) in their study. Burkart *et al.* (2018) have not mentioned clearly the number of controls they have used for COPD data analysis. Four studies have not mentioned the standard error of their effect size for lung function measures (Hancock *et al.* 2010; Wilk *et al.* 2012; Hansel *et al.* 2013; Lutz *et al.* 2015). There are even fewer GWAS on longitudinal lung function measures, which leaves a huge research gap to be filled. Loth *et al.* (2014) asserts that using cross-sectional measures of FVC made it difficult to determine whether the identified signals are due to influence on lung growth or age related lung function decline, thus emphasizing on the need for studies using longitudinal measures.

The identification of the genetic loci with biologically plausible functions in different populations warrants investigations in the future to elucidate the mechanisms underlying the effects they exert on pulmonary function (Hancock *et al.* 2010). Understanding the genetic factors related to reduced lung function and COPD susceptibility could offer clues for drug target identification, risk prediction, and stratified prevention or treatment (Wain *et al.* 2017). Researchers are suggesting assessment of the concept of 'diseasome', a link between the cellular networks and phenotypic manifestations, to be able to develop a common language for future research (Guo *et al.* 2012).

With a very few validation studies available, an urgent need for research is felt, in highly diverse populations in which only a limited number of genetic studies related to COPD have been performed using small sample sizes. Future meta-studies of lung function decline should aim to increase sample size while maintaining high phenotypic comparability among participating studies.

In conclusion, genomewide approaches have facilitated the detection of several genetic variants associated with lung volume in the Western population groups. These have thus increased our understanding of their underlying genetic architecture and presented us with data regarding the biological pathways related to the COPD development and lung function decline. The effect sizes of the variants in the genetic loci associated with lung function and COPD clarify only a modest proportion of the genetic variance. Racial differences in COPD may be present, which could be considered during sample selection and analysis. Understanding the genetic determinants of pulmonary function is crucial for the identification of the biological mechanisms that trigger its decline and ultimately reduce the mortality burden associated with the reduced pulmonary function. Elucidating the mechanisms involved in the COPD pathogenesis will enable a deeper understanding of lung function regulations and facilitate the development of new therapeutic targets for COPD.

Acknowledgment

This work was funded by University of Delhi's Research and Development Grant.

References

Artigas M. S., Loth D. W., Wain L. V., Gharib S. A., Obeidat M., Tang W. *et al.* 2011a Genome-wide association and large scale follow-up identifies 16 new loci influencing lung function. *Nat. Genet.* **43**, 1082–1090.

Artigas M. S., Wain L. V., Repapi E., Obeidat M., Sayers I., Burton P. I. *et al.* 2011b Effect of five genetic variants associated with lung function on the risk of chronic obstructive lung disease, and their joint effects on lung function. *Am. J. Respir. Cell Mol. Biol.* **184**, 786–795.

Artigas M. S., Wain L. V., Miller S., Kheirallah A. K., Huffman J. E., Ntalla I. *et al.* 2015 Sixteen new lung function signals identified through 1000 Genomes Project reference panel imputation. *Nat. Commun.* **6**, 8658.

Baines K. J., Fu J. J., McDonald V. M. and Gibson P. G. 2017 Airway gene expression of IL-1 pathway mediators predicts exacerbation risk in obstructive airway disease. *Int. J. Chron. Obstruct. Pulmon. Dis.* **12**, 541–550.

Boone T. 2014 Pulmonary ventilation. In *Introduction to exercise physiology* (ed. M. Jhonson, S. Coombs and J. Lundeen), pp. 17–38. , Jones & Bartlett Learning, Burlington.

Brandsma C., van den Berge M., Postma D. S., Jonker M. R., Brouwer S., Parè P. D. *et al.* 2015 A large lung gene expression study identifying fibulin-5 as a novel player in tissue repair in COPD. *Thorax* **70**, 21–32.

Burkart K. M., Sofer T., London S. J., Manichaikul A., Hartwig F. P., Yan Q. *et al.* 2018 A genome-wide association study in Hispanics/Latinos identifies novel signals for lung function. *Am. J. Respir. Crit. Care Med.* **198**, 208–219.

Chang Y., Glass K., Liu Y. Y., Silverman E. K., Crapo J. D., Tal-Singer R. *et al.* 2016 COPD subtypes identified by network-based clustering of blood gene expression. *Genomics* **107**, 51–58.

Chen W., Brehm J. M., Manichaikul A., Cho M. H., Boutaoui N., Yan Q. *et al.* 2015 A genome-wide association study of chronic obstructive pulmonary disease in Hispanics. *Ann. Am. Thorac. Soc.* **12**, 340–348.

Cho M. H., Boutaoui N., Klanderman B. J., Sylvia J. S., Ziniti J. P., Hersh C. P. *et al.* 2010 Variants in *FAM13A* are associated with chronic obstructive pulmonary disease. *Nat. Genet.* **42**, 200–202.

Cho M. H., Castaldi P. J., Wan E. S., Siedlinski M., Hersh C. P., Demeo D. L. *et al.* 2012 A genome-wide association study of COPD identifies a susceptibility locus on chromosome 19q13. *Hum. Mol. Genet.* **21**, 947–957.

Cho M. H., McDonald M. N., Zhou X., Mattheisen M., Castaldi P. J., Hersh C. P. *et al.* 2014 Risk loci for chronic obstructive pulmonary disease: a genome wide association study and meta-analysis. *Lancet Respir. Med.* **2**, 214–225.

Córdoba-Lanús E., Baz-Dávila R., Espinoza-Jiménez A., Rodríguez-Pérez M. C., Varo N., de-Torres J. P. *et al.* 2015

IL-8 gene variants are associated with lung function decline and multidimensional BODE index in COPD patients but not with disease susceptibility: a validation study. *Int. J. Chron. Obstruct. Pulmon. Dis.* **12**, 55–61.

Curjuric I., Imboden M., Nadif R., Kumar A., Schindler C., Huan M. *et al.* 2012 Different genes interact with particulate matter and tobacco smoke exposure in affecting lung function decline in the general population. *PLoS One* **7**, e40175.

Dijkstra A. E., Postma D. S., van Ginneken B., Wielpütz M. O., Schmidt M., Becker N. *et al.* 2015 Novel genes for airway wall thickness identified with combined genome-wide association and expression snalyses. *Am. J. Respir. Crit. Care Med.* **191**, 547–556.

Ding Y., Yang D., Zhou L., Xu J., Chen Y., He P. *et al.* 2015 Variants in multiple genes polymorphism association analysis of COPD in the Chinese Li population. *Int. J. Chron. Obstruct. Pulmon. Dis.* **10**, 1455–1463.

Eichler E. E., Flint J., Gibson G., Kong A., Leal S. M., Moore J. H. *et al.* 2010 Missing heritability and strategies for finding the underlying causes of complex disease. *Nat. Rev. Genet.* **11**, 446–450.

Farahani P., Halabian R., Vahedi E. and Salimian J. 2017 Increased genes expression levels of cytokines related to Th17/Treg cells in peripheral blood mononuclear cell correlate with clinical severity in COPD and mustard gas-exposed patients. *Iran J. Allergy Asthma Immunol.* **16**, 396–403.

Gibson G. 2010 Hints of hidden heritability in GWAS. *Nat. Genet.* **42**, 558–560.

Gibson G. 2011 Rare and common variants: twenty arguments. *Nat. Rev. Genet.* **13**, 135–145.

Guo Y., Gong Y., Pan C., Qian Y., Shi G., Cheng Q. *et al.* 2012 Association of genetic polymorphisms with chronic obstructive pulmonary disease in Chinese Han population: a case–control study. *BMC Med. Genomics* **5**, 64.

Han M. K., Agusti A., Calverley P., Celli B. R., Criner G., Curtis J. L. *et al.* 2010 Chronic obstructive pulmonary disease phenotypes: the future of COPD. *Am. J. Respir. Crit. Care Med.* **182**, 598–604.

Hancock D. B., Eijgelsheim M., Wilk J. B., Gharib S. A., Loehr L. R., Marciante K. R. *et al.* 2010 Meta-analyses of genome-wide association studies identify multiple novel loci associated with pulmonary function. *Nat. Genet.* **42**, 45–52.

Hancock D. B., Artigas M. S., Gharib S. A., Henry A., Manichaikul A., Ramasamy A. *et al.* 2012 Genome-wide joint meta-analysis of SNP and SNP-by-smoking interaction identifies novel loci for pulmonary function. *PLoS Genet.* **8**, e1003098.

Hansel N. N., Ruczinski I., Rafaels N., Sin D. D., Daley D., Malinina A. *et al.* 2013 Genome-wide study identifies two loci associated with lung function decline in mild to moderate COPD. *Hum. Genet.* **132**, 79–90.

Hardin M., Zielinski J., Wan E. S., Hersh C. P., Castaldi P. J., Schwinderr E. *et al.* 2012 CHRNA3/5, IREB2, and ADCY2 are associated with severe chronic obstructive pulmonary disease in Poland. *Am. J. Respir. Cell Mol. Biol.* **47**, 203–208.

Hernández-Montoya J., Pérez-Ramos J., Montaño M., Ramírez-Venegas A., Sansores R. H., Pérez- Rubio G. *et al.* 2015 Genetic polymorphisms of matrix metalloproteinases and protein levels in chronic obstructive pulmonary disease in a Mexican population. *Biomark. Med.* **9**, 979–988.

Hobbs B. D., de Jong K., Lamontagne M., Bosse Y., Shrine N., Artigas M. S. *et al.* 2017 Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nat. Genet.* **49**, 426–432.

Hole D. J., Watt G. C., Davey-Smith G., Hart C. L., Gillis C. R. and Hawthorne V. M. 1996 Impaired lung function and mortality risk in men and women: findings from the Renfrew and Paisley prospective population study. *BMJ* **313**, 711–716.

Hosgood III H. D., Minashe I., He X., Chanok S. and Lan Q. 2009 PTEN identified as important risk factor of chronic obstructive pulmonary disease. *Respir. Med.* **103**, 1866–1870.

Kato R., Mizuno S., Kadowaki M., Shiozaki K., Akai M., Nakawaga K. *et al.* 2016 Sirt1 expression is associated with CD31 expression in blood cells from patients with chronic obstructive pulmonary disease. *Respir. Res.* **17**, **139**.

Kim E. J., Kim K. M., Park S. H., Kim J. S., Lee W. K., Cha S. I. *et al.* 2008 SLC11A1 polymorphisms are associated with the risk of chronic obstructive pulmonary disease in a Korean population. *Biochem. Genet.* **46**, 506–519.

Lee B. Y., Cho S., Shin D. H. and Kim H. 2011 Genome-wide association study of copy number variations associated with pulmonary function measures in Korea associated resource (KARE) cohorts. *Genomics* **97**, 101–105.

Li X., Ortega V. E., Ampleford E. J., Barr R. G., Christenson S. A., Cooper C. B. *et al.* 2018 Genome-wide association study of lung function and clinical implication in heavy smokers. *BMC Med. Genet.* **19**, 134.

Liao S. Y., Lin X. and Christiani D. C. 2013 Gene-environment interaction effects on lung function- a genome-wide association study within the Framingham heart study. *Environ. Health* **12**, 101.

Liao S. Y., Lin X. and Christiani D. C. 2014 Genome-wide association and network analysis of lung function in the Framingham Heart Study. *Genet. Epidemiol.* **38**, 572–578.

Liu Y., Liang W., Gao L., Wang Y. and Zhang L. 2009 Association of CD40 - 1C/T polymorphism in the 5'-untranslated region and chronic obstructive pulmonary disease. *Clin. Chim. Acta* **408**, 56–59.

Loth D. W., Artigas M. S., Gharib S. A., Wain L. V., Franceschini N., Koch B. *et al.* 2014 Genome-wide association analysis identifies six new loci associated with forced vital capacity. *Nat. Genet.* **46**, 669–677.

Lutz S. M., Cho M. H., Young K., Hersh C. P., Castaldi P. J., McDonald M. L. *et al.* 2015 A genome-wide association study identifies risk loci for spirometric measures among smokers of European and African ancestry. *BMC Genet.* **16**, 138.

MacRedmond R. E., Greene C. M., Dorscheid D. R., McElvaney N. G. and O'Neill S. J. 2007 Epithelial expression of TLR4 is modulated in COPD and by steroids, salmeterol and cigarette smoke. *Respir. Res.* **8**, 84.

Manolio T. A., Collins F. S., Cox N. J., Goldstein D. B., Hindorff L. A., Hunter D. J. *et al.* 2009 Finding the missing heritability of complex diseases. *Nature* **461**, 747–753.

Myint P. K., Luben R. N., Surtees P. G., Wainwright N. W., Welch A. A., Bingham S. A. *et al.* 2005 Respiratory function and self-reported functional health: EPIC-Norfolk population study. *Eur. Respir. J.* **26**, 494–502.

Naghvi M., Wang H., Lozano R., Davis A., Liang X., Zhou M. *et al.* 2015 Global, regional, and national age-sex specific all-cause and cause specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease study 2013. *Lancet* **385**, 117–171.

Nedeljkovic I., Lahousse L., Carnero-Montoro E., Faiz A., Vonk J. M., deJong K. *et al.* 2018 COPD GWAS variant at 19q13.2 in relation with DNA methylation and gene expression. *Hum. Mol. Genet.* **27**, 396–405.

Ong B. A., Li J., McDonough J. M., Wei Z., Kim C., Chiavacci R. *et al.* 2013 Gene network analysis in a pediatric cohort identifies novel lung function genes. *PLoS One* **8**, e72899.

Pillai S. G., Ge D., Zhu G., Kong X., Shianna K. V., Need A. C. *et al.* 2009 A genome-wide association study in chronic

obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet.* **5**, e1000421.

Rabe K. F., Hurd S., Anzueto A., Barnes P. J., Buist S. A., Calverly P. *et al*. 2007 Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am. J. Respir. Crit. Care Med.* **176**, 532–555.

Raghu G., Collard H. R., Egan J. J., Martinez F. J., Behr J., Brown K. K. *et al*. 2011 An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am. J. Respir. Crit. Care Med.* **183**, 788–824.

Repapi E., Sayers I., Wain L. V., Burton P. L., Jhonson T., Obeidat M. *et al*. 2010 Genome-wide association study identifies five loci associated with lung function. *Nat. Genet.* **42**, 36–44.

Schünemann H. J., Dorn J., Grant B. J., Winkelstein W. J. and 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.

Siedlinski M., Cho M. H., Bakke P., Gulsvik A., Lomas D. A., Anderson W. *et al*. 2011 Genome-wide association study of smoking behaviours in COPD patients. *Thorax* **66**, 894–902.

Silverman E. K., Spira A. and Paré P. D. 2009 Genetics and genomics of chronic obstructive pulmonary disease. *Proc. Am. Thorac. Soc.* **6**, 539–542.

Strachan D. P. 1992 Ventilatory function, height, and mortality among lifelong non-smokers. *J. Epidemiol. Commun. Health* **46**, 66–70.

Tang W., Kowgier M., Loth D. W., Artigas M. S., Bonnie R. J., Emily H. *et al*. 2014 Large-scale genome-wide association studies and meta-analyses of longitudinal change in adult lung function. *PLoS One* **9**, e100776.

Takizawa H., Tanaka M., Takami K., Ohtoshi T., Ito K., Satoh M. *et al*. 2001 Increased expression of transforming growth factorb1 in small airway epithelium from tobacco smokers and patients with chronic obstructive pulmonary disease (COPD). *Am. J. Respir. Crit. Care Med.* **163**, 1476–1483.

Thun G. A., Imboden M., Ferrarotti I., Kumar A., Obeidat M., Zorzetto M. *et al*. 2013 Casual and synthetic associations of variants in the *SERPINA* gene cluster with alpha1-antitrypsin serum levels. *PLoS Genet.* **9**, e1003585.

Tripathi P. M., Kant S., Yadav R. S., Kushwaha R. A. S., Prakash V., Rizvi S. H. M. *et al*. 2017 Expression of toll-like receptor 2 and 4 in peripheral blood neutrophil cells from patients with chronic obstructive pulmonary disease. *Oman Med. J.* **32**, 477–485.

Wain L. V., Shrine N., Artigas M. S., Erzurumluoglu M. A., Noyvert B., Bossini-Castillo L. *et al*. 2017 Genome wide association analyses for lung function and chronic obstructive pulmonary disease identify new loci and potential druggable targets. *Nat. Genet.* **49**, 416–425.

Wang B., Zhou H., Yang J., Xiao J., Liang B., Li D. *et al*. 2013a Association of *HHIP* polymorphisms with COPD and COPD-related phenotypes in a Chinese Han population. *Gene* **531**, 101–105.

Wang B., Liang B., Yang J., Xiao J., Ma C., Xu S. *et al*. 2013b Association of *FAM13A* polymorphisms with COPD and COPD-related phenotypes in a Han Chinese. *Clin. Biochem.* **46**, 1683–1688.

Wang H., Naghavi M., Allen C., Barber R. M., Bhutta Z. A., Carter A. *et al*. 2016 Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* **388**, 1459–1544.

Wang I. M., Stepaniants S., Boie Y., Mortimer J. R., Kennedy B., Elliott M. *et al*. 2008 Gene expression profiling in patients with chronic obstructive pulmonary disease and lung cancer. *Am. J. Respir. Crit. Care Med.* **177**, 402–411.

Wilk J. B., Djousse L., Arnett D. K., Rich S. S., Province M. A., Hunt S. C. *et al*. 2000 Evidence for major genes influencing pulmonary function in the NHLBI Family Heart Study. *Genet. Epidemiol.* **19**, 81–94.

Wilk J. B., Walter R. E., Laramie J. M., Gottlieb D. J. and O'Connor G. T. 2007 Framingham Heart study genome-wide association: results for pulmonary function measures. *BMC Med. Genet.* **8**, S8.

Wilk J. B., Chen T., Gottlieb D. J., Walter R. E., Nagle M. W., Brandler B. J. *et al*. 2009 A Genome-wide association study of pulmonary function measures in the framingham heart study. *PLoS Genet.* **5**, e1000429.

Wilk J. B., Shrine N. R. G., Loehr R. L., Zhao J. H., Manichaikul A., Lopez L. M. *et al*. 2012 Genome-wide association studies identify CHRNA5/3 and HTR4 in the development of airflow obstruction. *Am. J. Respir. Crit. Care Med.* **186**, 622–632.

Young R. P., Hopkins R. and Eaton T. E. 2007 Forced expiratory volume in one second: not just a lung function test but a marker of premature death from all causes. *Eur. Respir. J.* **30**, 616–622.

Yuan Y., Jiang H., Kuang J., Hou X., Feng Y. and Su Z. 2012 Genetic variations in *ADIPOQ* gene are associated with chronic obstructive pulmonary disease. *PLoS One* **7**, e50848.

Zappala C. J., Latsi P. I., Nicholson A. G., Colby T. V., Cramer D., Renzoni E. A. *et al*. 2010 Marginal decline in forced vital capacity is associated with a poor outcome in idiopathic pulmonary fibrosis. *Eur. Respir. J.* **35**, 830–836.

Zhou X., Baron R. M., Hardin M., Cho M. H., Zielinski J., Hawrylkiewicz I. *et al*. 2012 Identification of chronic obstructive pulmonary disease genetic determinant that regulates HHIP. *Hum. Mol. Genet.* **21**, 1325–1335.

Ziółkowska-Suchanek I., Mosor M., Gabryel P., Grabicki M., Żurawek M., Fichna M. *et al*. 2015 Susceptibility loci in lung cancer and COPD: association of IREB2 and FAM13A with pulmonary diseases. *Sci. Rep.* **5**, 13502.

Corresponding editor: Analabha Basu