The protective effect of club cell secretory protein (CC-16) on COPD risk and progression: a Mendelian randomisation study

SUPPLEMENTARY RESULTS

AUTHORS:

Stephen Milne[1,2,3]*, Xuan Li[1]*, Ana I Hernandez Cordero[1], Chen Xi Yang[1], Michael H Cho[4], Terri H Beaty[5], Ingo Ruczinski[6], Nadia N Hansel[7], Yohan Bossé[8], Corry-Anke Brandsma[9], Don D Sin[1,2], Ma'en Obeidat[1]
*equal contributions to the manuscript

1. Centre for Heart Lung Innovation, St Paul's Hospital and University of British Columbia, Vancouver, BC, Canada
2. Division of Respiratory Medicine, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada
3. Faculty of Medicine and Health, University of Sydney, Sydney, New South Wales, Australia
4. Channing Division of Network Medicine and Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA, USA
5. Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA
6. Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA
7. Pulmonary and Critical Care Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, USA
8. Institut universitaire de cardiologie et de pneumologie de Québec, Department of Molecular Medicine, Laval University, Quebec City, Canada
9. University of Groningen Department of Pathology and Medical Biology, University Medical Centre Groningen, Groningen, The Netherlands

CORRESPONDING AUTHOR:

Dr Stephen Milne
UBC Centre for Heart Lung Innovation
Rm 166, St Paul's Hospital
1081 Burrard Street,
Vancouver, BC, V6Z 1Y6
CANADA

E: Stephen.milne@hli.ubc.ca

T: +1 604 806 8346

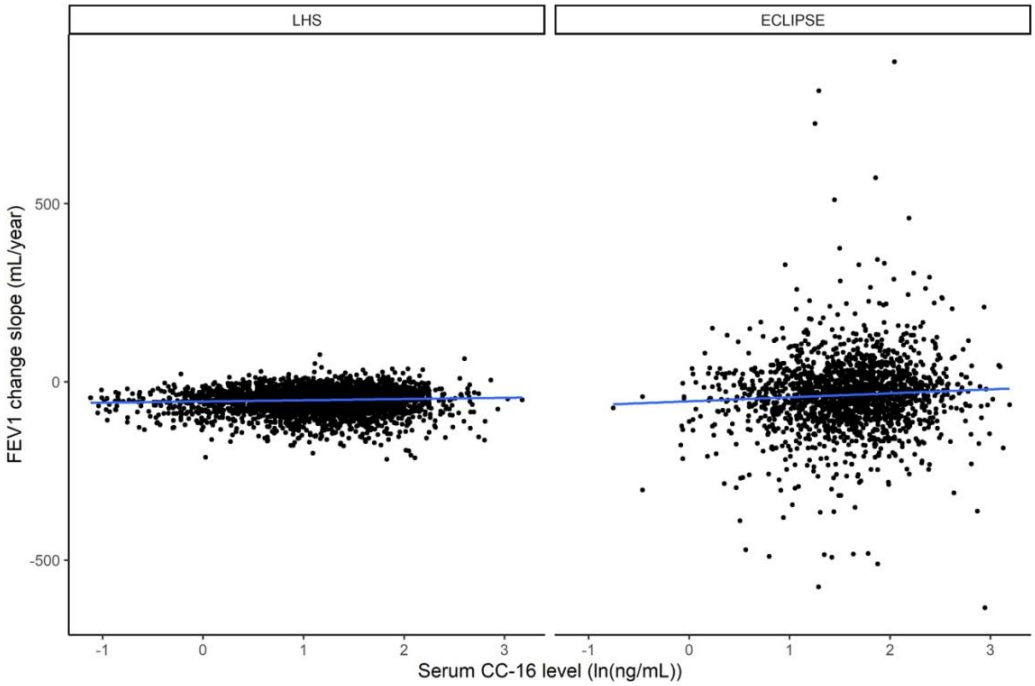
CONTENTS

1. TABLE S1: ASSOCIATION BETWEEN SERUM CC-16 LEVEL AND CHANGE IN FEV_1 IN THE BIOMARKER COHORTS

	n	CC-16 effect on annual change in FEV1 (β)†	SE	p
LHS	3444	2.43	0.93	0.01*
ECLIPSE	1821	6.25	3.94	0.11
Meta-analysis	5265	2.64	0.91	0.004**

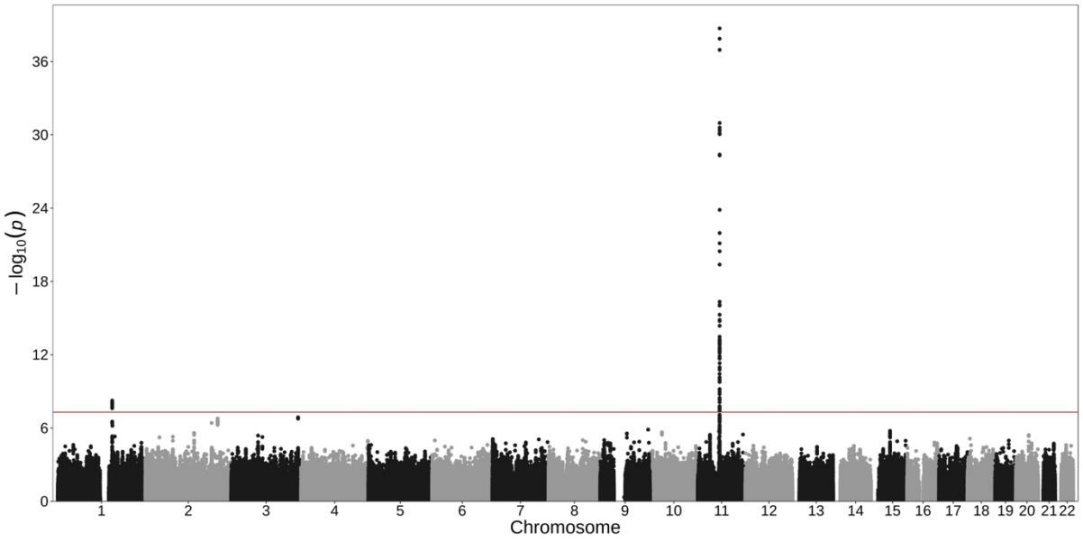
Multiple linear mixed effects model for change in FEV_1 with ln(CC-16), adjusted for age, sex, smoking status, baseline forced expiratory volume in 1 s (FEV1), body mass index, and the interactions of each factor with time. †change in annual rate of FEV1 decline (mL/year) per unit increase in ln(ng/mL) CC-16. SE: standard error. LHS: Lung Health Study. *p<0.05 **p<0.01.

2. FIGURE S1: ANNUAL CHANGE IN FEV_1 VERSUS SERUM CC-16 CONCENTRATION



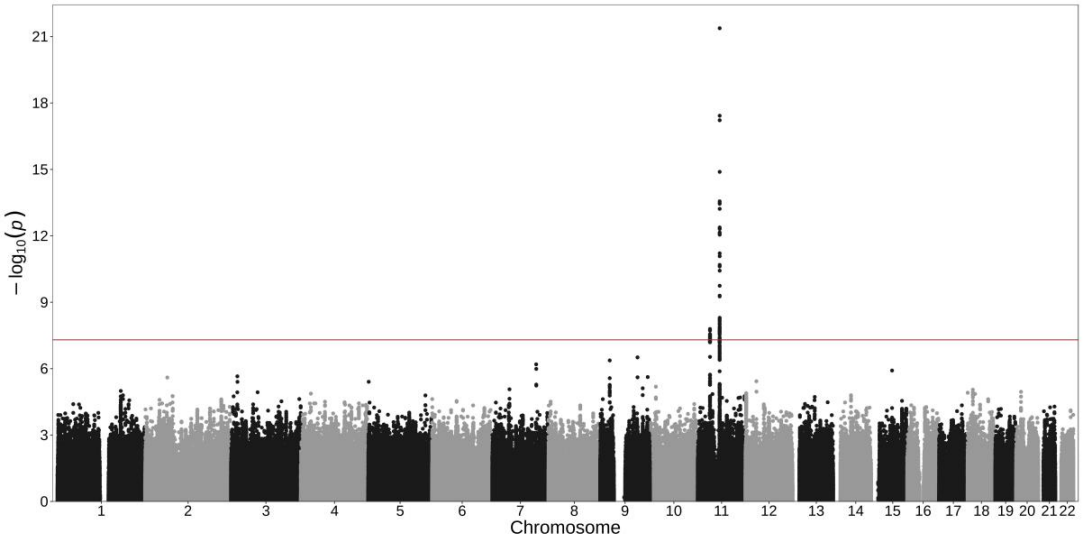
Annual change in FEV1 (calculated as the slope of FEV_1 versus time) versus CC-16 concentration in ln(ng/mL). Blue line = robust linear regression (unadjusted).

3. FIGURE S2: MANHATTAN PLOT FOR SERUM CC-16 GENOME-WIDE ASSOCIATION STUDY (GWAS) IN LUNG HEALTH STUDY



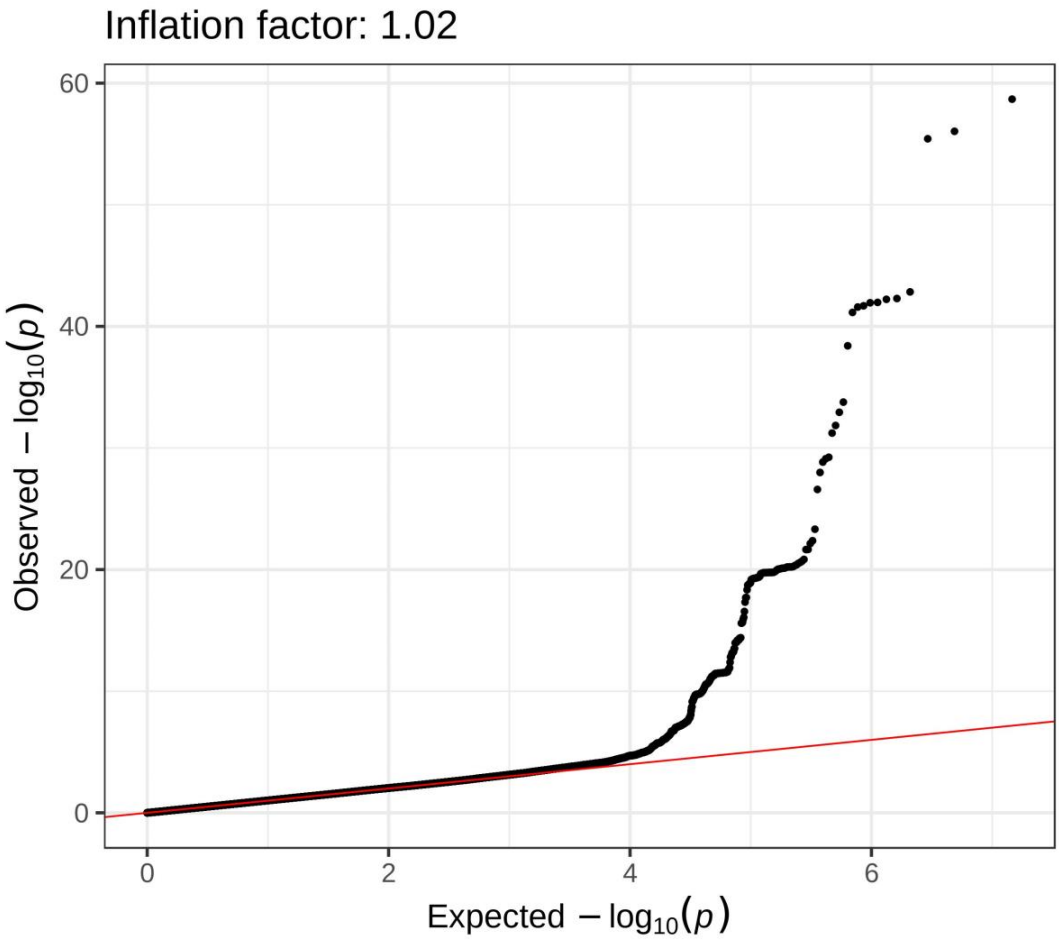
GWAS p values (–$\log_{10}$ scale) (Y axis) versus single nucleotide polymorphism positions across 22 chromosomes (X axis). Horizontal red line represents the genome-wide significance cut-off of $5x10^{-8}$. CC-16, club cell secretory protein-16.

4. FIGURE S3: MANHATTAN PLOT FOR SERUM CC-16 GENOME-WIDE ASSOCIATION STUDY (GWAS) IN ECLIPSE STUDY



GWAS p values (–$\log_{10}$ scale) (Y axis) versus single nucleotide polymorphism positions across 22 chromosomes (X axis). Horizontal red line represents the genome-wide significance cut-off of $5x10^{-8}$. CC-16, club cell secretory protein-16.

5. FIGURE S4: QUANTILE-QUANTILE PLOT FOR SERUM CC-16 GENOME-WIDE ASSOCIATION STUDY (GWAS)



Meta-analysis of Lung Health Study and ECLIPSE study GWAS. CC-16, club cell secretory protein-16. Observed p-values ($-\log_{10}$ scale) (Y axis) versus expected p-values ($-\log_{10}$ scale) (X axis). Red line represents where observed p values are equal to the expected.

6. TABLE S2: JOINT AND CONDITIONAL ANALYSES ON CHROMOSOME 11 CC-16 PROTEIN QUANTITATIVE TRAIT LOCI (pQTLS)

				Joint analysis			Conditional analysis		
SNP rsID	**Chr**	**Position§**	**Effect/Alt allele**	**Allele effect β†**	**SE**	**p**	**Allele effect β†**	**SE**	**p**
rs11032840	11	34779464	G/T	0.076	0.011	1.79×10^{-12}	0.076	0.011	1.79×10^{-12}
rs3741240	11	62186542	G/A	0.117	0.013	9.91×10^{-19}	-	-	-
rs11231085	11	62190448	G/C	0.110	0.013	3.37×10^{-16}	0.077	0.011	8.59×10^{-12}

Joint analysis: estimated joint effects of the three pQTLs on chromosome 11. Conditional analysis: estimated effect of the pQTLs on serum CC16 levels conditional on the top pQTL (rs3741240). §hg19 build of human reference genome. †change in ln(ng/mL) CC-16 per effect allele. SNP: single nucleotide polymorphism. rsID: reference SNP cluster identifier. Chr: chromosome. SE: standard error.

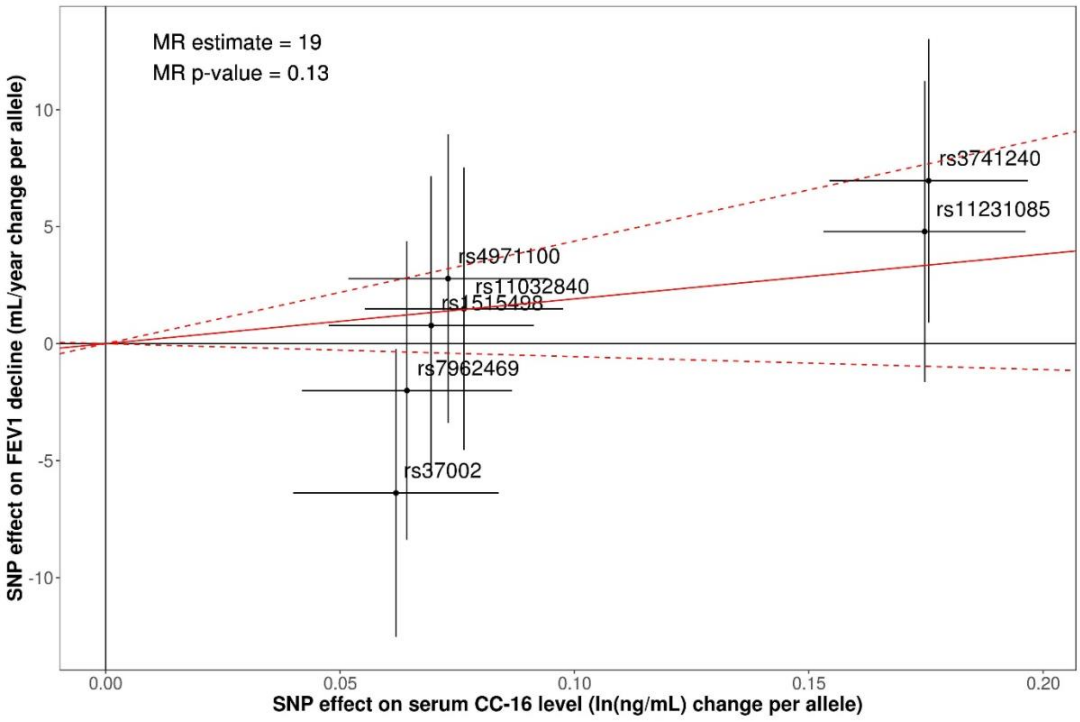
7. TABLE S3: ASSOCIATIONS BETWEEN CC-16 PROTEIN QUANTITATIVE TRAIT LOCI (pQTLS) AND COPD OUTCOMES

				COPD risk			COPD progression (Change in FEV_1 in mL/year)					
				ICGC			LHS			ECLIPSE		
SNP rsID	Chr	Position§	Effect/Alt allele	Allele effect β†	SE	p	Allele effect β‡	SE	p	Allele effect β	SE	p
rs4971100	1	155155731	A/G	-3.9×10^{-3}	0.01	0.7	-1.06	0.81	0.19	2.78	3.15	0.38
rs1515498	3	189508302	A/G	-2.2×10^{-3}	0.01	0.83	1.43	0.83	0.08	0.78	3.26	0.81
rs37002	5	1356944	C/T	-0.02	0.01	0.13	-0.57	0.85	0.50	-6.38	3.14	0.04*
rs11032840	11	34779464	G/T	3.8×10^{-3}	0.01	0.71	0.10	0.79	0.90	1.49	3.08	0.63
rs3741240	11	62186542	G/A	-0.02	0.01	0.14	1.63	0.82	0.05*	6.96	3.09	0.02*
rs11231085	11	62190448	G/C	-0.02	0.01	0.11	2.89	0.83	5.4×10^{-4}*	4.79	3.28	0.14
rs7962469	12	52348259	A/G	-0.03	0.01	1.2×10^{-3}*	-0.51	0.87	0.56	-2.00	3.25	0.54

§hg19 build of human reference genome. †ln(odds ratio) for COPD per effect allele CC-16 per effect allele. ‡change in annual rate of FEV_1 decline (mL/year) per effect allele *p<0.05. ICGC, International COPD Genetics Consortium; COPD, chronic obstructive pulmonary disease; LHS, Lung Health Study; ECLIPSE, Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints; SNP, single nucleotide polymorphism; rsID, reference SNP cluster identifier; Chr, chromosome; Alt, alternate; SE, standard error.

8. FIGURE S5: IVW MENDELIAN RANDOMIZATION (MR) PLOT FOR COPD PROGRESSION IN ECLIPSE STUDY

Inverse variance weighted regression model, adjusted for linkage disequilibrium between single-nucleotide polymorphisms (SNPs), intercept constrained to zero. The model relates the per-allele effects of the SNPs on serum CC-16 level to their per-allele effects on lung function decline. The red line represents the estimated effect. Error bars represent 95% confidence intervals. SNPs are annotated by their rs identifier. CC-16, club cell secretory protein-16; FEV_1, forced expiratory volume in 1 second.



9. TABLE S4: GWAS CATALOGUE LOOK-UP RESULTS

rsID	Chr	Effect allele	Trait	P value	Sample	Reference
rs4971100	1	A	Magnesium levels	$1x10^{-07}$	2,317 European, 1,283 African American ancestry children	Chang et al[1]
rs4971100	1	A	Magnesium levels	$4x10^{-07}$	2,317 European, 1,283 African American ancestry children	Chang et al[1]
rs4971100	1	A	Estimated glomerular filtration rate	$9x10^{-17}$	567,460 European, 165,726 East Asian, 13,842 African American, 13,359 South Asian, 4,961 Hispanic ancestry individuals	Wuttke et al[2]
rs4971100	1	A	Estimated glomerular filtration rate	$2x10^{-07}$	165,726 East Asian ancestry individuals	Wuttke et al[2]
rs4971100	1	A	Serum uric acid levels	$8x10^{-19}$	121,745 Japanese ancestry individuals, at least 88,461 European ancestry individuals	Nakatochi et al[3]
rs4971100	1	A	Serum uric acid levels	$5x10^{-16}$	121,745 Japanese ancestry individuals, at least 88,461 European ancestry individuals	Nakatochi et al[3]
rs3741240	11	A	Chronic obstructive pulmonary disease-related biomarkers	$1x10^{-26}$	Up to 1,951 European ancestry smokers	Kim et al[4]
rs7962469	12	?	Lung function (FEV1/FVC)	$8x10^{-19}$	Approximately 370,000 European ancestry individuals	Kichaev et al[5]

Taken from the NHGRI-EBI Catalog of Human Genome-Wide Association Studies (https://www.ebi.ac.uk/gwas/) which lists known SNP-trait associations with $p<1x10^{-5}$.

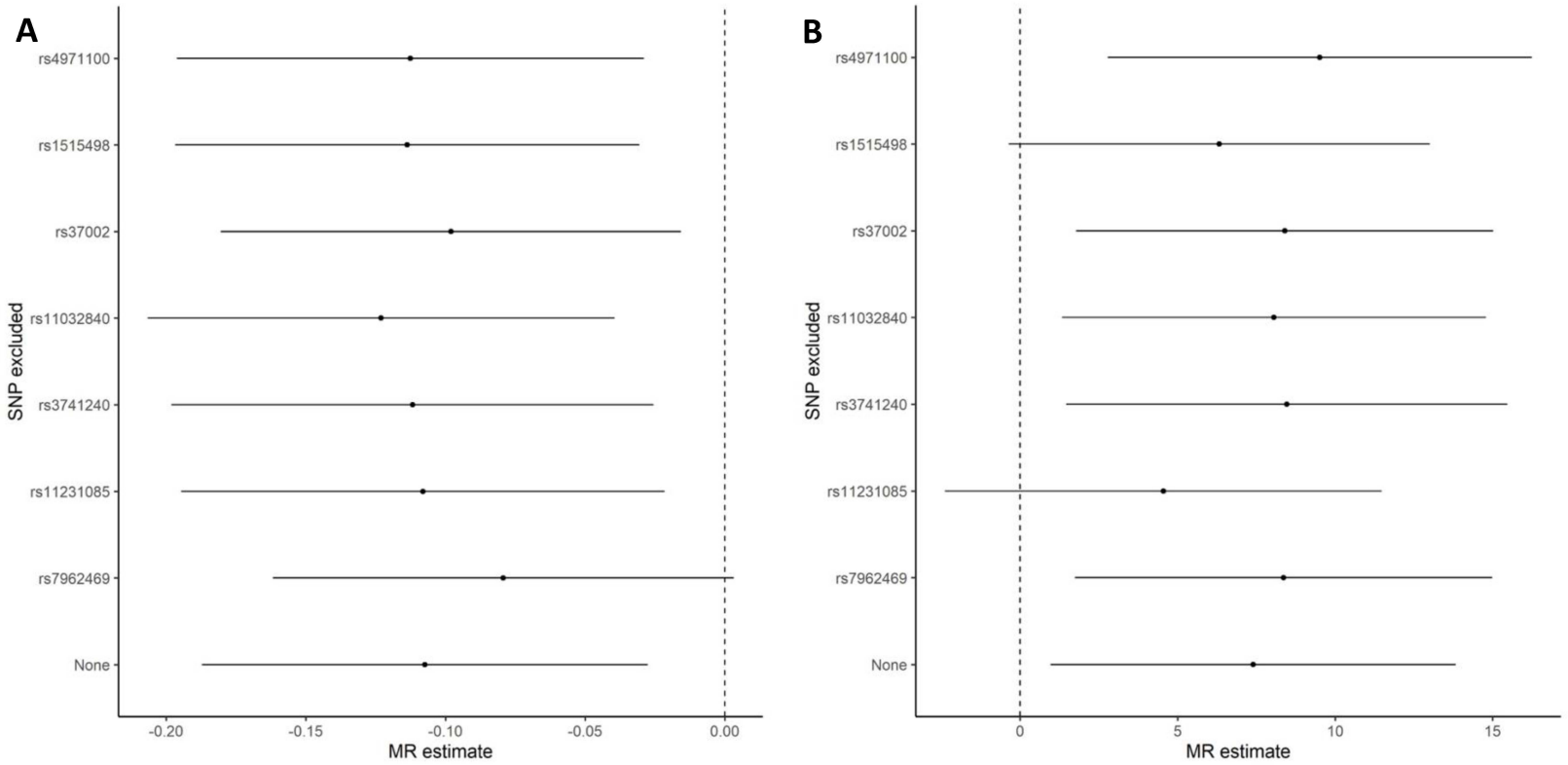
10.TABLE S5: PARTIAL F STATISTICS FOR INDIVIDUAL INTRUMENTAL VARIABLES

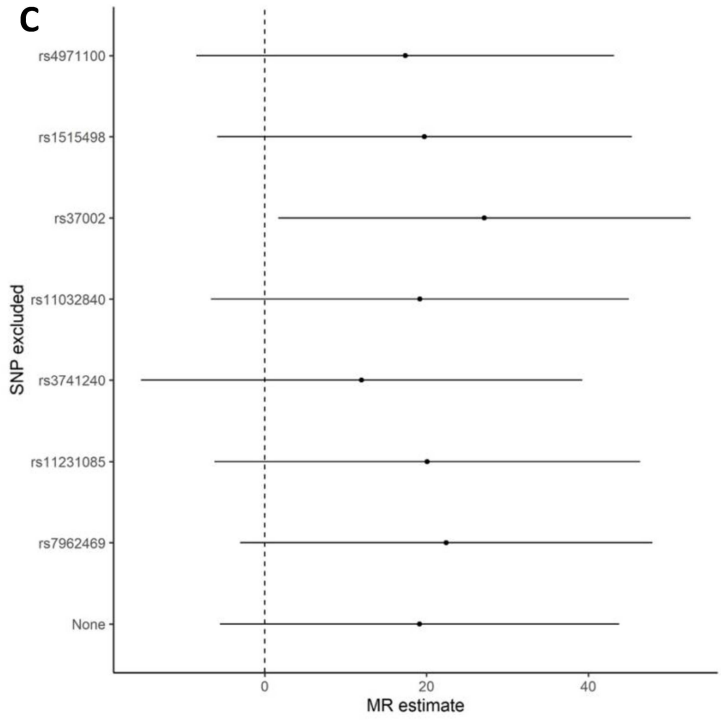
		Partial F	
rsID	**Chr**	**LHS**	**ECLIPSE**
rs4971100	1	34.09	11.93
rs1515498	3	27.86	11.14
rs37002	5	11.52	21.45
rs11032840	11	21.47	31.17
rs3741240	11	168.23	96.09
rs11231085	11	176.63	76.2
rs7962469	12	16.97	14.78

F statistics derived from ANOVA comparing linear models for ln(CC-16) concentration with and without the SNP as covariate. rsID, reference SNP cluster identifier; Chr, chromosome.

11.FIGURE S6: LEAVE-ONE-OUT SENSITIVITY ANALYSES

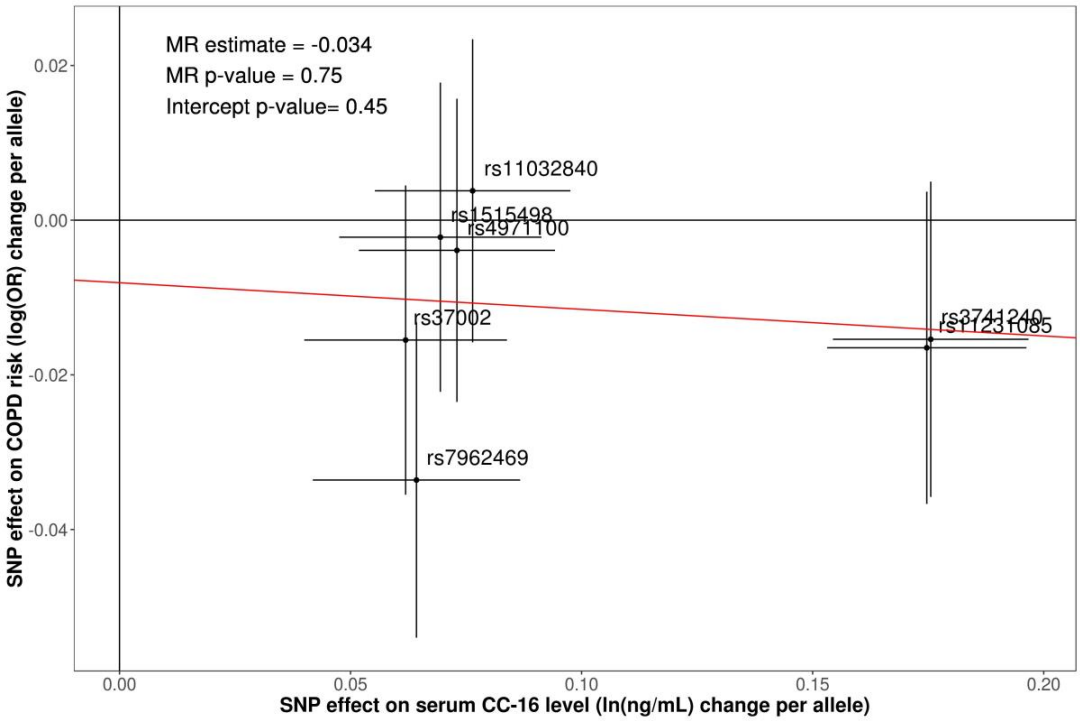
Mendelian randomisation (MR) estimates and 95% confidence intervals from inverse variance weighted models using all 7 SNPs ("None") and after removing a single SNP (as indicated on Y axis). MR outcomes: (A) "COPD risk" in ICGC dataset, (B) "COPD progression" (annual change in FEV_1) in Lung Health Study, (C) "COPD progression" in ECLIPSE study.




C
rs4971100
rs1515498
rs37002
rs11032840
rs3741240
rs11231085
rs7962469
None
SNP excluded
0
20
40
MR estimate

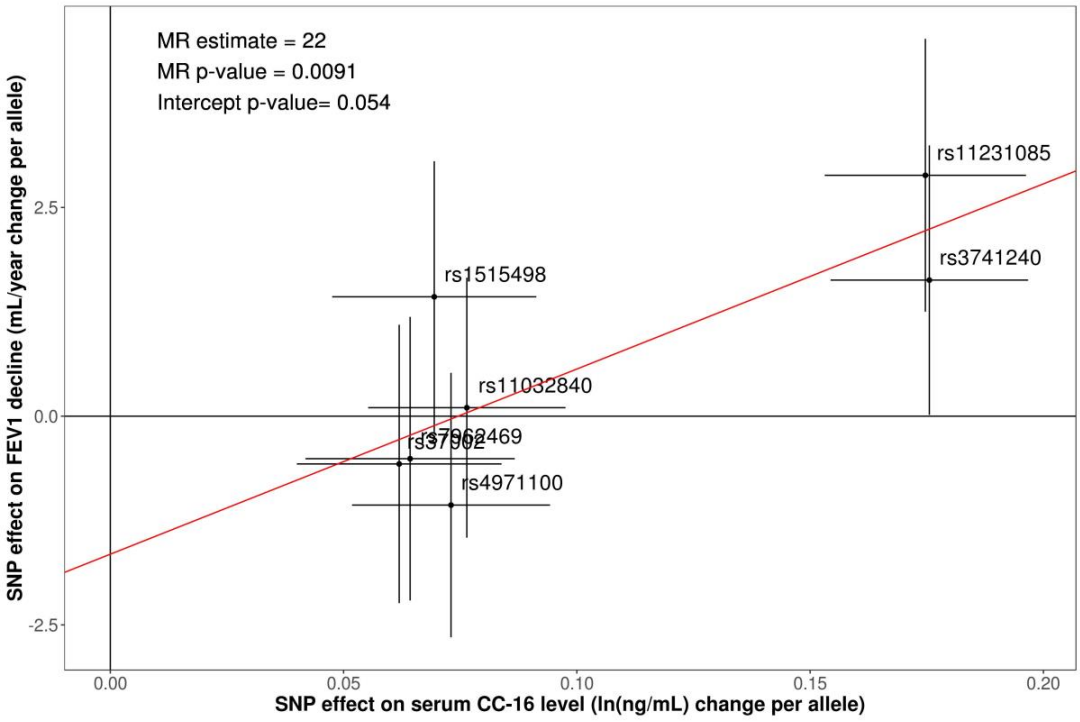
12.FIGURE S7: MENDELIAN RANDOMISATION (MR) EGGER ANALYSIS FOR COPD RISK

Inverse variance weighted regression model, adjusted for linkage disequilibrium between single-nucleotide polymorphisms (SNPs), with unconstrained intercept. The model relates the per-allele effects of the SNPs on serum CC-16 level to their per-allele effects on risk of having chronic obstructive pulmonary disease (COPD) in the International COPD Genetics Consortium (ICGC) dataset. The red line represents the estimated effect. Error bars represent 95% confidence intervals. SNPs are annotated by their rs identifier. OR, odds ratio; CC-16, club cell secretory protein-16.

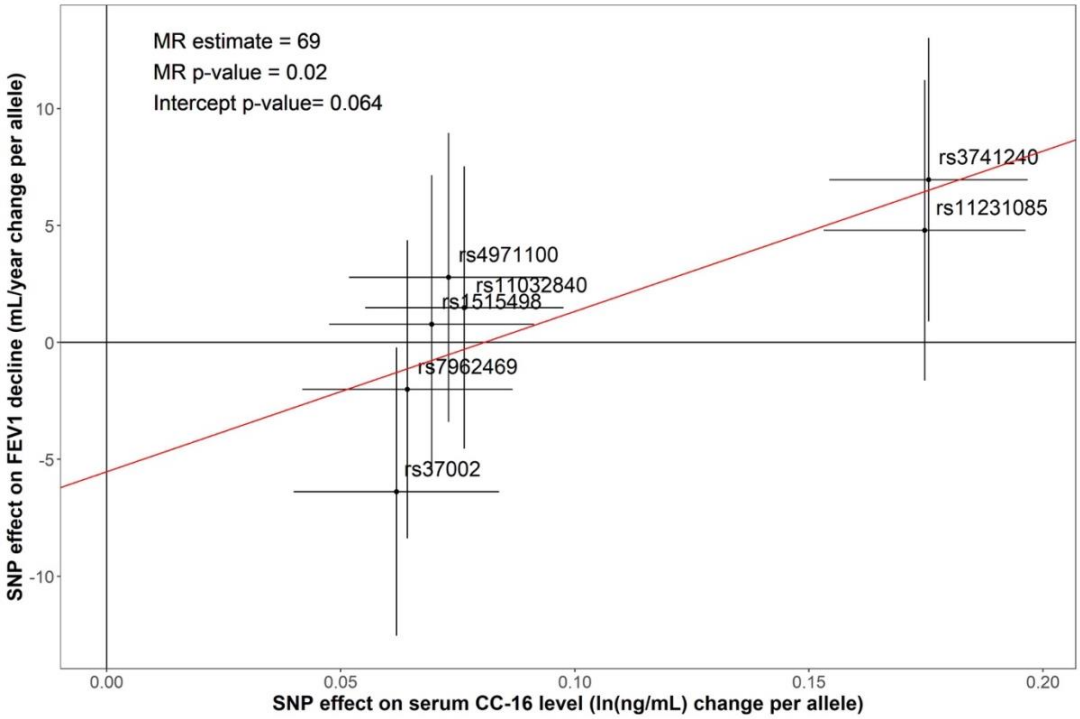


13. FIGURE S8: MENDELIAN RANDOMISATION (MR) EGGER ANALYSIS FOR COPD PROGRESSION IN THE LUNG HEALTH STUDY

Inverse variance weighted regression model, adjusted for linkage disequilibrium between single-nucleotide polymorphisms (SNPs), with unconstrained intercept. The model relates the per-allele effects of the SNPs on serum CC-16 level to their per-allele effects on lung function decline. The red line represents the estimated effect. Error bars represent 95% confidence intervals. SNPs are annotated by their rs identifier. CC-16, club cell secretory protein-16, FEV_1, forced expiratory volume in 1 second.



14. FIGURE S9: MENDELIAN RANDOMISATION (MR) EGGER ANALYSIS FOR COPD PROGRESSION IN THE ECLIPSE STUDY



Inverse variance weighted regression model, adjusted for linkage disequilibrium between single-nucleotide polymorphisms (SNPs), with unconstrained intercept. The model relates the per-allele effects of the SNPs on serum CC-16 level to their per-allele effects on lung function decline. The red line represents the estimated effect. Error bars represent 95% confidence intervals. SNPs are annotated by their rs identifier. CC-16, club cell secretory protein-16, FEV_1, forced expiratory volume in 1 second.

15.REFERENCES

1. Chang X, Li J, Guo Y, et al. Genome-wide association study of serum minerals levels in children of different ethnic background. *PLoS ONE* 2015;10(4):e0123499. doi: 10.1371/journal.pone.0123499 [published Online First: 2015/04/18]
2. Wuttke M, Li Y, Li M, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat Genet* 2019;51(6):957-72. doi: 10.1038/s41588-019-0407-x [published Online First: 2019/06/04]
3. Nakatochi M, Kanai M, Nakayama A, et al. Genome-wide meta-analysis identifies multiple novel loci associated with serum uric acid levels in Japanese individuals. *Communications biology* 2019;2:115. doi: 10.1038/s42003-019-0339-0 [published Online First: 2019/04/18]
4. Kim DK, Cho MH, Hersh CP, et al. Genome-wide association analysis of blood biomarkers in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2012;186(12):1238-47. doi: 10.1164/rccm.201206-1013OC [published Online First: 2012/11/13]
5. Kichaev G, Bhatia G, Loh PR, et al. Leveraging polygenic functional enrichment to improve GWAS power. *Am J Hum Genet* 2019;104(1):65-75. doi: 10.1016/j.ajhg.2018.11.008 [published Online First: 2019/01/01]