Genetic and epigenetic associations with pre-COPD lung function trajectories

David J Martino[1], Dinh S Bui[2] Shuai Li[2,3,4,5], Sabrina Idrose[2], Jennifer Perret[2], Adrian Lowe[2], Caroline Lodge[2], Gayan Bowatte[2], Yuben Moodley[6], Paul S Thomas[7], Graeme Zosky[8], Philip M Hansbro[9], John Holloway[10], Cecile Svanse [11,12], Rosa Faner[13,14,15], Eugene H Walters [2,16*] and Shaymali C Dharmage [2] *.

** Joint Senior authors*

1 Wal-yan Respiratory Research Centre, Telethon Kids Institute, University of Western Australia.

2 Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne

3 Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge

4 Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Victoria, Australia

5 Murdoch Children's Research Institute, Royal Children's Hospital, Parkville, Victoria 3051, Australia

6 Department of Respiratory Medicine, Fiona Stanley Hospital, Murdoch, Australia

7 Respiratory Medicine Prince of Wales' Hospital & Prince of Wales' Clinical School, UNSW, Randwick, NSW 2031, Australia

8 Menzies Institute for Medical Research, University of Tasmania

9 Centre for Inflammation, Centenary Institute and University of Technology Sydney, School of Life Sciences, Faculty of Science, Sydney NSW 2007, Australia

10 Human Development and Health, Faculty of Medicine, University of Southampton, Southampton UK.

11 Centre for International Health, Department of Global Public Health and Primary Care, University of Bergen, Norway.

12 Department of Occupational Medicine, Haukeland University Hospital, Bergen, Norway.

13 University of Barcelona, Immunology Unit, Biomedicine Department. Spain.

AJRCCM Articles in Press. Published August 23, 2023 as 10.1164/rccm.202306-1025LE
Copyright © 2023 by the American Thoracic Society

14. IDIBAPS-FCRB Institut d’Investigacions Biomediques August Pi i Sunyer-Fundació Clinic Recerca, Barcelona, Spain.

15. Centro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBER), Spain.

16 School of Medicine, University of Tasmania and School of Population and Global Health, University of Melbourne

Running title: Epigenetics of pre-COPD lung function trajectories.

Corresponding author: David Martino

Telethon Kids Institute

PO Box 855, West Perth, Western Australia, 6872

David.Martino@telethonkids.org.au

Abbreviations

COPD – Chronic Obstructive Pulmonary Disease

CpG – Cytosine phosphate Guanine

DMR – Differentially Methylated Region

GWAS – Genome Wide Association Study

mQTL – Methylation Quantitative Trait Loci

SNP – Single Nucleotide Polymorphism

TAHS – Tasmanian Longitudinal Health Study

Author contributions: SD, DJM, DB participated in the conception and design of the study.

RF, SL, NS, JP, EHW, SI participated in the intellectual development, interpretation of data

Copyright © 2023 by the American Thoracic Society

and drafting of the manuscript. YM, GZ, PMH, JH, CS, AL, CL contributed to the development and editing of the manuscript.

AJRCCM Articles in Press. Published August 23, 2023 as 10.1164/rccm.202306-1025LE
Copyright © 2023 by the American Thoracic Society

To the Editor:

Understanding the molecular mechanisms of lung function trajectories that progress to chronic obstructive pulmonary disease (COPD) (pre-COPD trajectories) will inform preventive interventions. The Tasmanian Longitudinal Health Study (TAHS) previously defined lifetime FEV_1 (forced expiratory volume in one second) trajectories using serial spirometry measures conducted in a cohort from age seven to 53 years (1). Of the six pre-bronchodilator FEV_1 lifetime trajectories identified, three collectively accounted for 75% of chronic obstructive pulmonary disease (COPD) prevalence at age 53 years (2). These high-risk trajectories were: 1) early below average lung function (usual rate of subsequent decline), 2) persistently low, and 3) early below average lung function with accelerated decline. The TAHS cohort provides a unique opportunity to investigate molecular factors associated with these high-risk trajectories and we conducted a mechanistic pilot study to inform more extensive longitudinal studies in the future.

Epigenetic modifications such as DNA methylation are established in early embryonic development and incorporate variation according to host genotype and modifiable factors across the lifespan. Previous large-scale studies have demonstrated associations between DNA methylation markers and pulmonary function (3) and COPD in both blood and lung tissue (4, 5). Declining lung function has been associated with metastable epigenetic associations in blood (6), as well as epigenetic age acceleration (proxy markers of accelerated biological age) (6, 7). We conducted a screen of both epigenetic (DNA methylation) and genetic (single nucleotide polymorphisms; SNPs) markers on available whole blood samples collected at the 45-year follow-up. We used an extremes of phenotype design to maximise power randomly selecting 80 subjects from across the three high-risk trajectories (labelled as high-risk

Copyright © 2023 by the American Thoracic Society

trajectory group) and matching on age and smoking status to 80 subjects from the persistently high trajectory (control group).

By design, individuals belonging to the different lung function trajectories varied in post bronchodilator FEV_1, but also steroid medication use, and sex so these differences were examined in adjusted models. We quantified 787,111 DNA methylation markers (CpGs) and 4,456,571 SNPs using the InfiniumMethylationEPIC (v1) and Infinium Global Screening Array (v3) genotyping microarrays. Some of the results of this study have been previously reported in the form of a conference abstract (8).

We detected DNA methylation changes at 55 differentially methylated regions (DMRs) containing 73 unique genes and 6 non-coding regions (FDR adjusted $P < 0.05$; Figure 1A) between the control group and high-risk trajectory group. Notable affected genes included *LY6G5C and HLA-DQB1* within the major histocompatibility complex, HOX cluster transcription factors (*HOXB-AS3, HOXB3, HOXB6)* that have been implicated in the pathogenesis of pulmonary diseases (9, 10), and transmembrane glycoprotein genes (*LGALS3BP, OCA2, KCNE1, PTPRN2, TNXB, PCDHGA5, CDSN, PCDHGA4, PCDHGA3, PCDHGB3, PCDHGA2, PCDHGB2, PCDHGA1, EGFR, DPP6, FOLH1, SGCD, CRTAC1, PCDHGB1, FIBIN, CHST1, MUC4, DPEP3)* that play a role in epithelial biology including in the formation of epithelial to mesenchymal transition (11).

This DMR signature was only partially consistent across the three high-risk trajectory sub-groups (24% of DMRs shared across all sub-groups), whereas sub-group specific regions predominated (Figure 1B) consistent with the notion that different COPD risk-factor etiotypes are to likely exhibit different molecular drivers (12). Current COPD or current asthma explained 17-30% of methylation differences across the DMRs respectively, but sex and blood cell counts were not mediators or confounders of these associations. By integrating the

AJRCCM Articles in Press. Published August 23, 2023 as 10.1164/rccm.202306-1025LE
Copyright © 2023 by the American Thoracic Society

genetic and epigenetic data sets and performing methylation quantitative trait (mQTL) mapping we found that genetic variation at 381 nearby SNPs (+/- 500kb of DMRs) in 17 genomic loci were associated with 23% of CpGs within DMR regions (Figure 1C). The strongest mQTL region was on chromosome 6 at the major histocompatibility locus. Using publicly available tissue specific gene expression signatures (GTEx catalogue v8) we determined these mQTL SNPs were most strongly associated with transcripts primarily expressed in the lung (*ATP13A4, MUC4, PSORS1C1*) (Figure 1D). Several mQTLs have previously been associated with spirometry outcomes (*HAPLN1, HLA-DRB1, HLA-DQA1, HLA-DQB1*), COPD (*HLA-DQB1, HLA-DQA1*) and asthma risk (*HLA-DQB1, HLA-DQA1, HLA-DRB1, HLA-DRB6, PSORS1C1*) in the genome-wide association study (GWAS) catalogue. Consistent with previous studies (6, 7) we also found that epigenetic age acceleration was significantly higher in the combined high-risk trajectory group when measured using the phenoAge algorithm (13) (Table 1). On average, individuals in the high-risk trajectory group had increased mean predicted chronological age of 1.5 years relative to controls (40.2 v 38.7, P=0.03, t-test). Stratified analysis suggested age-acceleration was strongest in the early below average, accelerated decline trajectory (Beta = 2.1, P=0.06 v 1.4, P=0.11, *below average*; v 1.1, P=0.19, *persistently low*) although not significant, probably related to the small sample size.

To our knowledge this was the first epigenome-wide association analysis in individuals from COPD high-risk lung function trajectories, providing a strong foundation for further delineation of phenotypes and risk factors to enable precision molecular profiling. We determined blood to be a phenotypically relevant tissue to explore molecular associations with life-time lung function trajectories in this cohort. Although causality of the epigenetic associations cannot be established in this pilot study, a subset of epigenetic changes in the high-risk trajectory were mQTLs whereby allelic variation influenced the methylation patterns

AJRCCM Articles in Press. Published August 23, 2023 as 10.1164/rccm.202306-1025LE
Copyright © 2023 by the American Thoracic Society

at nearby genes. Since the causality of genetic variation on DNA methylation levels is uni-directional, these analyses aid in prioritizing methylation-trait associations from epigenome-wide scans. The mQTL associations are especially compelling candidates for gene-environment interactions and might prove to be linked to early life events, as well as processes related to disease progression. Confirmatory longitudinal studies are now planned to dissect these potential environmental and host genomic risk factors that are reflected in the epigenome. Our finding of low DMR sharing across sub-groups of the high-risk trajectory group suggests molecular risk factors are unique across different life-time lung function trajectories warranting follow-up studies at cohort-wide scale. We also found that respiratory morbidities including current asthma and COPD explained a proportion of variation in the blood epigenetic markers, and in agreement with other cohorts (6, 7) epigenetic age acceleration was detectable in the high-risk trajectory group, independently of COPD or asthma status.

In summary this pilot study confirms the utility of our approach and paves the way for future epigenetic profiling studies in this unique cohort including extending this pilot study to the full cohort and investigating profiles of lifetime spirometry patters (14). An enhanced understanding of molecular mechanisms associated with spirometry patterns such as obstruction and restriction will enable more precise biomarker-driven interventions in the future with potential to redirect the course of respiratory health in those with pre-COPD.

AJRCCM Articles in Press. Published August 23, 2023 as 10.1164/rccm.202306-1025LE
Copyright © 2023 by the American Thoracic Society

References

1. Matheson MC, Abramson MJ, Allen K, Benke G, Burgess JA, Dowty JG, Erbas B, Feather IH, Frith PA, Giles GG, Gurrin LC, Hamilton GS, Hopper JL, James AL, Jenkins MA, Johns DP, Lodge CJ, Lowe AJ, Markos J, Morrison SC, Perret JL, Southey MC, Thomas PS, Thompson BR, Wood-Baker R, Haydn Walters E, Dharmage SC, group Ti. Cohort Profile: The Tasmanian Longitudinal Health STUDY (TAHS). *Int J Epidemiol* 2017; 46: 407-408i.
2. Bui DS, Lodge CJ, Burgess JA, Lowe AJ, Perret J, Bui MQ, Bowatte G, Gurrin L, Johns DP, Thompson BR, Hamilton GS, Frith PA, James AL, Thomas PS, Jarvis D, Svanes C, Russell M, Morrison SC, Feather I, Allen KJ, Wood-Baker R, Hopper J, Giles GG, Abramson MJ, Walters EH, Matheson MC, Dharmage SC. Childhood predictors of lung function trajectories and future COPD risk: a prospective cohort study from the first to the sixth decade of life. *Lancet Respir Medicine* 2018; 6: 535-544.
3. Lee M, Huan T, McCartney DL, Chittoor G, de Vries M, Lahousse L, Nguyen JN, Brody JA, Castillo-Fernandez J, Terzikhan N, Qi C, Joehanes R, Min JL, Smilnak GJ, Shaw JR, Yang CX, Colicino E, Hoang TT, Bermingham ML, Xu H, Justice AE, Xu CJ, Rich SS, Cox SR, Vonk JM, Prokic I, Sotoodehnia N, Tsai PC, Schwartz JD, Leung JM, Sikdar S, Walker RM, Harris SE, van der Plaat DA, Van Den Berg DJ, Bartz TM, Spector TD, Vokonas PS, Marioni RE, Taylor AM, Liu Y, Barr RG, Lange LA, Baccarelli AA, Obeidat M, Fornage M, Wang T, Ward JM, Motsinger-Reif AA, Hemani G, Koppelman GH, Bell JT, Gharib SA, Brusselle G, Boezen HM, North KE, Levy D, Evans KL, Dupuis J, Breeze CE, Manichaikul A, London SJ. Pulmonary Function and Blood DNA Methylation: A Multiancestry Epigenome-Wide Association Meta-analysis. *Am J Respir Crit Care Med* 2022; 206: 321-336.
4. Morrow JD, Glass K, Cho MH, Hersh CP, Pinto-Plata V, Celli B, Marchetti N, Criner G, Bueno R, Washko G, Choi AMK, Quackenbush J, Silverman EK, DeMeo DL. Human Lung DNA Methylation Quantitative Trait Loci Colocalize with Chronic Obstructive Pulmonary Disease Genome-Wide Association Loci. *Am J Respir Crit Care Med* 2018; 197: 1275-1284.
5. Qiu W, Baccarelli A, Carey VJ, Boutaoui N, Bacherman H, Klanderman B, Rennard S, Agusti A, Anderson W, Lomas DA, DeMeo DL. Variable DNA methylation is associated with chronic obstructive pulmonary disease and lung function. *Am J Respir Crit Care Med* 2012; 185: 373-381.
6. Carmona JJ, Barfield RT, Panni T, Nwanaji-Enwerem JC, Just AC, Hutchinson JN, Colicino E, Karrasch S, Wahl S, Kunze S, Jafari N, Zheng Y, Hou L, DeMeo DL, Litonjua AA, Vokonas PS, Peters A, Lin X, Schwartz J, Schulz H, Baccarelli AA. Metastable DNA methylation sites associated with longitudinal lung function decline and aging in humans: an epigenome-wide study in the NAS and KORA cohorts. *Epigenetics* 2018; 13: 1039-1055.
7. Rezwan FI, Imboden M, Amaral AFS, Wielscher M, Jeong A, Triebner K, Real FG, Jarvelin M-R, Jarvis D, Probst-Hensch NM, Holloway JW. Association of adult lung function with accelerated biological aging. *Aging Albany Ny* 2020; 12: 518-542.
8. Martino D, Bui D, Lodge C, Perret J, Boweatte G, Lowe A, Faner R, Walters E, Thomas P, Moodley Y, Zosky G, Hamilton G, Holloway J, Svanes C, Abramson M, Dharmage S. Epigenetic study of COPD-risk lung function trajectories. *Respirology* 2022; 27: 3-4.
9. Golpon HA, Geraci MW, Moore MD, Miller HL, Miller GJ, Tuder RM, Voelkel NF. HOX genes in human lung: altered expression in primary pulmonary hypertension and emphysema. *Am J Pathol* 2001; 158: 955-966.

AJRCCM Articles in Press. Published August 23, 2023 as 10.1164/rccm.202306-1025LE
Copyright © 2023 by the American Thoracic Society

10. Portas L, Pereira M, Shaheen SO, Wyss AB, London SJ, Burney PGJ, Hind M, Dean CH, Minelli C. Lung Development Genes and Adult Lung Function. *Am J Respir Crit Care Med* 2020; 202: 853-865.
11. Sohal SS, Mahmood MQ, Walters EH. Clinical significance of epithelial mesenchymal transition (EMT) in chronic obstructive pulmonary disease (COPD): potential target for prevention of airway fibrosis and lung cancer. *Clin Transl Med* 2014; 3: 33.
12. Stolz D, Mkorombindo T, Schumann DM, Agusti A, Ash SY, Bafadhel M, Bai C, Chalmers JD, Criner GJ, Dharmage SC, Franssen FME, Frey U, Han M, Hansel NN, Hawkins NM, Kalhan R, Konigshoff M, Ko FW, Parekh TM, Powell P, Mölken MR-v, Simpson J, Sin DD, Song Y, Suki B, Troosters T, Washko GR, Welte T, Dransfield MT. Towards the elimination of chronic obstructive pulmonary disease: a Lancet Commission. *Lancet* 2022; 400: 921-972.
13. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, Hou L, Baccarelli AA, Stewart JD, Li Y, Whitsel EA, Wilson JG, Reiner AP, Aviv A, Lohman K, Liu Y, Ferrucci L, Horvath S. An epigenetic biomarker of aging for lifespan and healthspan. *Aging Albany Ny* 2018; 10: 573-591.
14. Dharmage SC, Bui DS, Walters EH, Lowe AJ, Thompson B, Bowatte G, Thomas P, Garcia-Aymerich J, Jarvis D, Hamilton GS, Johns DP, Frith P, Senaratna CV, Idrose NS, Wood-Baker RR, Hopper J, Gurrin L, Erbas B, Washko GR, Faner R, Agusti A, Abramson MJ, Lodge CJ, Perret JL. Lifetime spirometry patterns of obstruction and restriction, and their risk factors and outcomes: a prospective cohort study. *Lancet Respir Med* 2023; 11: 273-282.

AJRCCM Articles in Press. Published August 23, 2023 as 10.1164/rccm.202306-1025LE
Copyright © 2023 by the American Thoracic Society

Figure Legends

Figure 1 – (A) Volcanoplot of differentially methylated regions**.** Each point represents a genomic region, and the number of individual CpGs in the region is shown on the y-axis as a function of the effect size (x-axis) interpreted as the percent change in methylation ratios (10^{-2}). (**B**) Upset plot showing the number of overlapping DMRs per trajectory sub-group. *ph*= *persistently high*; *acc.dec = accelerated decline; bl.ave = below average.* **(C)** Boxplot of the Mucin 4 mQTL showing methylation ratios expressed as a percentage (10^{-2}) stratified by genotype. Means comparisons by t-test, exact P-values shown. **(D)** Summary statistics of tissue-specific enrichment testing for mQTLs and sets of differentially expressed genes for 30 general tissue types in the GTExv8 catalogue.

AJRCCM Articles in Press. Published August 23, 2023 as 10.1164/rccm.202306-1025LE
Copyright © 2023 by the American Thoracic Society

Tables

Table 1 - Logistic regression of PhenoAge clock with case - control group

	Coefficient	Std. Error	z value	P value
AA	0.09	0.05	1.85	0.06
Male sex	1.11	0.38	2.90	0.06
Asthma	2.23	0.48	4.67	<0.01*
EAA	0.09	0.05	1.85	0.06
Male sex	1.11	0.38	2.90	<0.01*
Asthma	2.23	0.48	4.66	<0.01*
IEAA	0.10	0.05	1.94	0.05*
Male sex	1.09	0.38	2.87	<0.01*
Asthma	2.26	0.48	4.70	<0.01*

*Outcome variable = high-risk/persistently low, predictors: AA = Age acceleration residual, EAA = Extrinsic age acceleration residual, IEAA = Intrinsic age acceleration residual. * = P <0.05.*

AJRCCM Articles in Press. Published August 23, 2023 as 10.1164/rccm.202306-1025LE
Copyright © 2023 by the American Thoracic Society

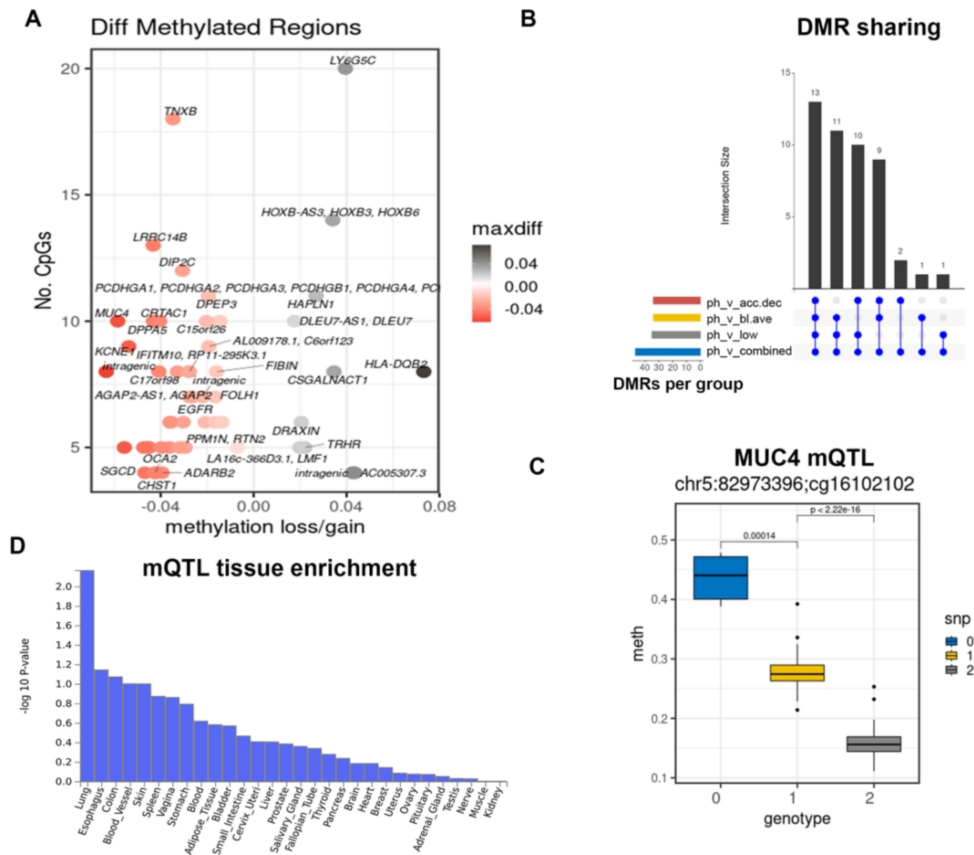

Figure 1 – (A) Volcanoplot of differentially methylated regions. Each point represents a genomic region, and the number of individual CpGs in the region is shown on the y-axis as a function of the effect size (x-axis) interpreted as the percent change in methylation ratios (10-2). (B) Upset plot showing the number of overlapping DMRs per trajectory sub-group. ph= persistently high; acc.dec = accelerated decline; bl.ave = below average. (C) Boxplot of the Mucin 4 mQTL showing methylation ratios expressed as a percentage (10-2) stratified by genotype. Means comparisons by t-test, exact P-values shown. (D) Summary statistics of tissue-specific enrichment testing for mQTLs and sets of differentially expressed genes for 30 general tissue types in the GTExv8 catalogue.

104x93mm (300 x 300 DPI)

AJRCCM Articles in Press. Published August 23, 2023 as 10.1164/rccm.202306-1025LE
Copyright © 2023 by the American Thoracic Society