HHS Public Access

Author manuscript

*Eur Respir J*. Author manuscript; available in PMC 2018 March 27.

Published in final edited form as:

*Eur Respir J*. 2017 November ; 50(5): . doi:10.1183/13993003.02042-2017.

# Genetic studies as a tool for identifying novel potential targets for treatment of COPD

**Ani Manichaikul**[1] and **Jennifer N. Nguyen**[1]

[1]Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

Chronic Obstructive Pulmonary Disease (COPD), characterized by chronic airflow limitation and abnormal response to noxious particles or gases [1] is currently the fourth leading cause of death worldwide [2] and the only major cause of death that has seen continued increases in recent years. While smoking cessation, oxygen therapy and bronchodilators offer some relief from the symptoms of disease, and medications are available to prevent COPD exacerbations [3], there are currently no treatments available to slow or stop progression, reflecting a limited understanding of the molecular underpinnings of the disease. In this issue of the *European Respiratory Journal*, Obeidat *et al.* [4] identify surfactant protein-D (SP-D) as a potentially causal risk factor for COPD. The paper is presented in multiple distinct parts, beginning with a genome-wide association study to identify SNPs from three genomic regions (on chromosomes 6, 10 and 16) associated with circulating levels of SP-D, followed by a Mendelian Randomization study demonstrating that genetically determined lower levels of SP-D are associated with increased risk of COPD based on results from a large-scale consortium based on a Genome-wide association study (GWAS) of up to ~11,000 COPD cases and ~37,000 controls recently published by the International COPD Genetics Consortium (ICGC) [5]. The results from the study of Obeidat *et al.* [4] are promising and open new potential avenues for early diagnosis and treatment of this disease. How did we get to this novel finding, what steps remain to carry this finding towards establishing the role of SP-D in COPD, and how can we stimulate additional discovery of novel therapeutic targets?

## Genetic studies of COPD support SP-D as a candidate biomarker

For decades, the *SERPINA1* gene that encodes α-1 antitrypsin was the only genetic factor known to accelerate decline in pulmonary function and increase risk of COPD. Replacement of α-1 antitrypsin may slow the decline for a subset of COPD patients with α-1 antitrypsin deficiency [6]. Over the past decade, GWAS by the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) [7] and SpiroMeta consortia [8] have identified ~30 gene regions that contain common single nucleotide polymorphisms (SNPs) significantly associated with Forced Expiratory Volume in 1 second ($FEV_1$) and its ratio to Forced Vital Capacity ($FEV_1$/FVC). GWAS approaches have also been used to identify genetic variants for COPD [5, 9–13]. Major candidate gene regions identified through GWAS of COPD

Correspondence: Ani Manichaikul, PhD, Assistant Professor of Public Health Sciences, Center for Public Health Genomics, University of Virginia School of Medicine, West Complex Room 6115, Charlottesville, VA 22903, amanicha@virginia.edu, Phone: 434-982-1612, Fax: 434-982-1815.

include *FAM13A* [11], *CHRNA3/5/IREB2* [9], and *HHIP* [13], with four novel loci, *EEFSEC*, *DSP*, *MTCL1*, and *SFTPD*, that emerged in the most recently published GWAS of COPD [5].

Before its identification through GWAS, SP-D had been already studied as a biomarker for COPD exacerbations [14] and asthma severity [15, 16]. Interestingly, the lead *SFTPD* SNP identified in association with COPD [5], rs721917, is a protein coding missense mutation providing a rationale for SP-D as the protein implicated by the identified genetic association. The same SNP has also demonstrated association with emphysema [17]. Despite the mounting evidence of SP-D as a biomarker for COPD, thus far there has not been sufficient evidence to establish a causal role for SP-D in the development of COPD.

## Mendelian randomization analysis and its underlying assumptions

Mendelian randomization provides a useful analytic framework to establish the causality of a candidate biomarker in pathogenesis of disease. The approach, which can be conducted in the context of an observational study, builds on the random assignment of alleles at meiosis to carry out what some consider a “natural” randomized trial. The basic idea underlying the approach is that we can use data from genetic association studies, including GWAS, to represent the biomarker of interest. In the study of Obeidat *et al.* [4], SNPs identified in GWAS of SP-D levels were used as instruments independent of potential confounders of the relationship between SP-D and COPD. When applied correctly, Mendelian randomization analysis can provide one piece of evidence to establish the biomarker of interest as a causal factor in the disease of interest [18].

Importantly, Mendelian Randomization analysis relies on multiple assumptions in order to make a valid inference regarding the causality of the underlying biomarker of interest [19, 20]. Here, we present the three underlying assumptions and discuss the extent to which these assumptions hold true in the context of the current study.

1. *Assumption 1: The SNPs used to construct the instrument should be robustly associated with SP-D.* The study of Obeidat *et al.* [4] used a stringent threshold for genome-wide significance to identify the SNPs taken forward for Mendelian Randomization analyses, such that we have relatively strong evidence that the assumption is met. However, as with any GWAS study, it is possible that some of the identified SNPs represent false positives. Although some of the pQTL reported in this study have been reported in previous studies [21, 22], some of the SP-D associated SNPs used in Mendelian Randomization analyses of Obeidat *et al.* [4] were newly reported in this study and have yet to be confirmed in analysis of independent data sets.

2. *Assumption 2: The SNPs used to construct the instrument should be unrelated to any external confounders of the relationship between SP-D and COPD.* SP-D has been associated with COPD even after individuals have ceased smoking [23], suggesting that the relationship between SP-D and COPD may be at least partially independent of this major risk factor for COPD. That said, it is difficult to provide concrete “proof” that this assumption is met, as there are potentially

additional confounders of the relationship between SP-D and COPD that have not yet been identified.

3. *Assumption 3: The SNPs used to construct the instrument for SP-D are related to COPD only through their modulation of SP-D levels.* While SP-D region SNPs have previously been reported in genetic association studies of COPD [5, 24] and emphysema [17], it is noteworthy that the top associated SNP in these studies was the protein coding missense mutation rs721917, providing a strong rationale to pinpoint SP-D as the underlying mediator. However, numerous other genes lie within ~200 kb of the *SFTPD* gene that encodes SP-D, among these *TMEM25A* encoding a transmembrane protein, the antisense RNA gene *NUTM2B-AS1*, and the mannose-binding lectin pseudogene, *MBL1P.* As the functionality of some GWAS-identified regions has been shown to act over regulatory regions that span over megabase distances [25], additional assessment of genes within the region of the SP-D locus may be needed to establish whether or not *SFTPD* is the lone gene underlying this genomic signal. That said, we agree with the authors that there is reasonable evidence to surmise SP-D might be the primary and possibly sole gene underlying the identified genomic signal at this locus.

## Mendelian randomization analysis: strengths, limitations and extensions

The core strength of the Mendelian randomization approach is that, applied correctly, it has the ability to inform on the causal nature of the relationship of the identified biomarker, in our case SP-D, and the disease outcome of interest, or COPD in this case. However, as noted above, there are multiple key assumptions necessary to make valid inference regarding causality based on the results of this Mendelian randomization study, and it remains somewhat unclear whether or not these assumptions held true for the current study. In addition, while the direction of effect provided by the reported Mendelian randomization analysis is informative in conceptualizing the general relationship between SP-D levels and risk of COPD, the corresponding effect estimates obtained by Mendelian randomization analysis likely do not generalize to the increased risk of COPD that may correspond to decreased levels of measured SP-D at the population level. Therefore, a successful Mendelian randomization study is merely the beginning, and in the future we would like to see (1) population-level longitudinal studies used to confirm the prognostic value of SP-D as a biomarker for COPD, and (2) randomized controlled clinical trials to assess the value of increasing SP-D for prevention of COPD.

Given the promising results obtained by the current Mendelian randomization study of SP-D for COPD, this study can also be viewed as proof of concept for characterization of other candidate biomarkers for COPD. The growing availability of biomarker studies for COPD present a good resource for future Mendelian randomization studies that could be used to assess the causal nature of those biomarkers in COPD. In addition, existing and forthcoming ‘omics studies provide the opportunity to apply the Mendelian randomization approach in a high-throughput manner. The methodologies for high-throughput extension of Mendelian randomization-type approaches are quickly gaining traction in the genetic community [26, 27] providing systematic frameworks for integration of genomic and transcriptomic data

sets. In the coming years, we expect it will become feasible and fruitful to leverage these new statistical methodologies, combined with growing genomic data sets for COPD and related traits, as well as lung-specific ‘omics studies to identify novel causal biomarkers for COPD, leading to novel therapeutic targets for treatment of COPD.

## Acknowledgments

This work was supported by funding from NIH R01HL131565.

## References

1. Gerald LB, Bailey WC. Global Initiative for Chronic Obstructive Lung Disease. J Cardiopulm Rehabil. 2002; 22:234–244. [PubMed: 12202842]
2. Lozano R, Naghavi M, Lim S, Adair T, AlMazroa MA, Anderson LM, Barker-Collo S, Bell PML, Bennett D, Bhalla K, Blyth F, Boufous S, Chen H. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012; 380:2095–2128. [PubMed: 23245604]
3. Wedzicha JA, Calverley PMA, Albert RK, Anzueto A, Criner GJ, Hurst JR, Miravitlles M, Papi A, Rabe KF, Rigau D, Sliwinski P, Tonia T, Vestbo J, Wilson KC, Krishnan JA. Prevention of COPD exacerbations: a European Respiratory Society/American Thoracic Society guideline. Eur Respir J. 2017; 50:1602265. [PubMed: 28889106]
4. Obeidat M, Li X, Zhou G. Surfactant Protein D is a Causal Risk Factor for Chronic Obstructive Pulmonary Disease: Results of Mendelian Randomization. Eur Respir J. 2017
5. Hobbs BD, de Jong K, Lamontagne M, Bossé Y, Shrine N, Artigas MS, Wain LV, Hall IP, Jackson VE, Wyss AB, London SJ, North KE, Franceschini N, Strachan DP, Beaty TH, Hokanson JE, Crapo JD, Castaldi PJ, Chase RP, Bartz TM, Heckbert SR, Psaty BM, Gharib SA, Zanen P, Lammers JW, Oudkerk M, Groen HJ, Locantore N, Tal-Singer R, Rennard SI, et al. Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. Nat Genet. 2017; 49:426–432. [PubMed: 28166215]
6. Teschler H. Long-term experience in the treatment of α1-antitrypsin deficiency: 25 years of augmentation therapy. Eur Respir Rev. 2015; 24:46–51. [PubMed: 25726554]
7. Psaty BM, O’Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, Uitterlinden AG, Harris TB, Witteman JCM, Boerwinkle E. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium design of prospective meta-analyses of genome-wide association studies from 5 Cohorts. Circ Cardiovasc Genet. 2009:73–80. [PubMed: 20031568]
8. Artigas MS, Loth DW, Wain LV, Gharib SA, Obeidat M, Tang W, Zhai G, Zhao JH, Smith AV, Huffman JE, Albrecht E, Jackson CM, Evans DM, Cadby G, Fornage M, Manichaikul A, Lopez LM, Johnson T, Aldrich MC, Aspelund T, Barroso I, Campbell H, Cassano PA, Couper DJ, Eiriksdottir G, Franceschini N, Garcia M, Gieger C, Gislason GK, Grkovic I, et al. Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. Nat Genet. 2011; 43:1082–1090. [PubMed: 21946350]
9. Wilk JB, Shrine NRG, Loehr LR, Zhao JH, Manichaikul A, Lopez LM, Smith AV, Heckbert SR, Smolonska J, Tang W, Loth DW, Curjuric I, Hui J, Cho MH, Latourelle JC, Henry AP, Aldrich M, Bakke P, Beaty TH, Bentley AR, Borecki IB, Brusselle GG, Burkart KM, Chen TH, Couper D, Crapo JD, Davies G, Dupuis J, Franceschini N, Gulsvik A, et al. Genome-wide association studies identify CHRNA5/3 and HTR4 in the development of airflow obstruction. Am J Respir Crit Care Med. 2012; 186:622–632. [PubMed: 22837378]
10. Pillai SG, Ge D, Zhu G, Kong X, Shianna KV, Need AC, Feng S, Hersh CP, Bakke P, Gulsvik A, Ruppert A, Carlsen KCL, Roses A, Anderson W, Rennard SI, Lomas DA, Silverman EK, Goldstein DB. A Genome-Wide association study in chronic obstructive pulmonary disease (COPD): Identification of two major susceptibility loci. PLoS Genet. 2009:5.
11. Cho MH, Boutaoui N, Klanderman BJ, Sylvia JS, Ziniti JP, Hersh CP, DeMeo DL, Hunninghake GM, Litonjua AA, Sparrow D, Lange C, Won S, Murphy JR, Beaty TH, Regan EA, Make BJ, Hokanson JE, Crapo JD, Kong X, Anderson WH, Tal-Singer R, Lomas DA, Bakke P, Gulsvik A,

Pillai SG, Silverman EK. Variants in FAM13A are associated with chronic obstructive pulmonary disease. Nat Genet. 2010; 42:200–202. [PubMed: 20173748]

12. Cho MH, Castaldi PJ, Wan ES, Siedlinski M, Hersh CP, Demeo DL, Himes BE, Sylvia JS, Klanderman BJ, Ziniti JP, Lange C, Litonjua AA, Sparrow D, Regan EA, Make BJ, Hokanson JE, Murray T, Hetmanski JB, Pillai SG, Kong X, Anderson WH, Tal-Singer R, Lomas DA, Coxson HO, Edwards LD, MacNee W, Vestbo J, Yates JC, Agusti A, Calverley PMA, et al. A genome-wide association study of COPD identifies a susceptibility locus on chromosome 19q13. Hum Mol Genet. 2012; 21:947–957. [PubMed: 22080838]
13. Cho MH, McDonald M-LN, Zhou X, Mattheisen M, Castaldi PJ, Hersh CP, DeMeo DL, Sylvia JS, Ziniti J, Laird NM, Lange C, Litonjua AA, Sparrow D, Casaburi R, Barr RG, Regan EA, Make BJ, Hokanson JE, Lutz S, Dudenkov TM, Farzadegan H, Hetmanski JB, Tal-Singer R, Lomas DA, Bakke P, Gulsvik A, Crapo JD, Silverman EK, Beaty TH. Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. Lancet Respir Med. 2014; 2:214–225. [PubMed: 24621683]
14. Lomas DA, Silverman EK, Edwards LD, Locantore NW, Miller BE, Horstman DH, Tal-Singer R. Serum surfactant protein D is steroid sensitive and associated with exacerbations of COPD. Eur Respir J. 2009; 34:95–102. [PubMed: 19164344]
15. Benfante A, Battaglia S, Paterno A, Spatafora M, Scichilone N. Serum surfactant protein d is a potential biomarker of severity of asthma. Am J Respir Crit Care Med. 2015:191. no pagination. [PubMed: 25932660]
16. Benfante A, Battaglia S, Principe S, Di Mitri C, Paternò A, Spatafora M, Scichilone N. Asthmatics with high levels of serum surfactant protein D have more severe disease. Eur Respir J. 2016:1864–1867. [PubMed: 26989101]
17. Ishii T, Hagiwara K, Kamio K, Ikeda S, Arai T, Mieno MN, Kumasaka T, Muramatsu M, Sawabe M, Gemma A, Kida K. Involvement of surfactant protein D in emphysema revealed by genetic association study. Eur J Hum Genet. 2012; 20:230–235. [PubMed: 21934714]
18. Weir D. Biosocial Surveys Biosoc Surv. 2008
19. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet. 2013
20. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological Challenges in Mendelian Randomization. Epidemiology. 2014; 25:427–435. [PubMed: 24681576]
21. Kim DK, Cho MH, Hersh CP, Lomas DA, Miller BE, Kong X, Bakke P, Gulsvik A, Agust?? A, Wouters E, Celli B, Coxson H, Vestbo J, MacNee W, Yates JC, Rennard S, Litonjua A, Qiu W, Beaty TH, Crapo JD, Riley JH, Tal-Singer R, Silverman EK. Genome-wide association analysis of blood biomarkers in chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2012; 186:1238–1247. [PubMed: 23144326]
22. Sun W, Kechris K, Jacobson S, Drummond MB, Hawkins GA, Yang J, Chen TH, Quibrera PM, Anderson W, Barr RG, Basta PV, Bleecker ER, Beaty T, Casaburi R, Castaldi P, Cho MH, Comellas A, Crapo JD, Criner G, Demeo D, Christenson SA, Couper DJ, Curtis JL, Doerschuk CM, Freeman CM, Gouskova NA, Han MLK, Hanania NA, Hansel NN, Hersh CP, et al. Common Genetic Polymorphisms Influence Blood Biomarker Measurements in COPD. PLoS Genet. 2016:12.
23. Sims MW, Tal-Singer RM, Kierstein S, Musani AI, Beers MF, Panettieri Ra, Haczku A. Chronic obstructive pulmonary disease and inhaled steroids alter surfactant protein D (SP-D) levels: a cross-sectional study. Respir Res. 2008; 9:13. [PubMed: 18226251]
24. Foreman MG, Kong X, DeMeo DL, Pillai SG, Hersh CP, Bakke P, Gulsvik A, Lomas DA, Litonjua AA, Shapiro SD, Tal-Singer R, Silverman EK. Polymorphisms in surfactant protein-D are associated with chronic obstructive pulmonary disease. Am J Respir Cell Mol Biol. 2011; 44:316–322. [PubMed: 20448057]
25. Smemo S, Tena JJ, Kim K-H, Gamazon ER, Sakabe NJ, Gómez-Marín C, Aneas I, Credidio FL, Sobreira DR, Wasserman NF, Lee JH, Puviindran V, Tam D, Shen M, Son JE, Vakili NA, Sung H-K, Naranjo S, Acemel RD, Manzanares M, Nagy A, Cox NJ, Hui C-C, Gomez-Skarmeta JL, Nóbrega MA. Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature. 2014; 507:371–375. [PubMed: 24646999]

26. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, Jansen R, de Geus EJC, Boomsma DI, Wright FA, Sullivan PF, Nikkola E, Alvarez M, Civelek M, Lusis AJ, Lehtimäki T, Raitoharju E, Kähönen M, Seppälä I, Raitakari OT, Kuusisto J, Laakso M, Price AL, Pajukanta P, Pasaniuc B. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet. 2016; 48:245–252. [PubMed: 26854917]

27. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, Eyler AE, Denny JC, Nicolae DL, Cox NJ, Im HK. A gene-based association method for mapping traits using reference transcriptome data. Nat Genet. 2015; 47:1091–1098. [PubMed: 26258848]

**Take home message**

Genetic studies are a useful tool for identification of novel causal biomarkers and therapeutic targets for COPD.