Differential lung tissue gene expression in males and females: implications for the susceptibility to develop COPD

To the Editor:

Several studies have shown that sexual dimorphisms exist with respect to the susceptibility to smoking and the clinical presentation of chronic obstructive pulmonary disease (COPD) [1, 2]. Females with COPD have a faster annual decline in their forced expiratory volume in 1 s (FEV1) than males, even with a similar number of pack-years smoking [3, 4]. In addition, females develop a different type of COPD to males with less severe emphysematous destruction, but thicker small airway walls [5]. Several factors may contribute to the observed sex differences, *e.g.* smoking behaviour or exposures to other environmental agents [6, 7].

In the present study, we aim to investigate which intrinsic biological factors may explain these observed sex differences in COPD. To this end, we analysed genome-wide gene expression data in resected lung tissue samples available from three academic sites (see Gene Expression Omnibus accession number GSE23546). Details of the subject enrolment, methods of RNA extraction and gene expression profiling has been described previously [8]. From the total cohort of 1111 lung tissue samples available, we only included those derived from 581 (327 male and 254 female) current and ex-smokers with and without COPD, the selection criteria were described previously [9]. Statistical analyses were performed in R, version 3.0.1 (www.r-project. org). We adjusted the complete dataset for the first five principal components calculated excluding X- and Y-chromosome-linked genes. This approach was taken to filter out unwanted variation due to unknown technical factors; in this way we made sure not to filter out sex-associated effects [9]. To analyse differential gene expression in terms of sex, we used a linear model adjusting for age, pack-years and smoking status (*i.e.* current- *versus* ex-smoking), and performed an inverse-variance meta-analysis across the three cohorts. A Benjamini–Hochberg False Discovery Rate (FDR) was applied to control for multiple testing and an FDR <0.05 was considered statistically significant. The independent replication cohort consisted of non-involved lung tissue samples derived from 202 males and 72 females who underwent lobectomy for adenocarcinoma at several hospitals near Milan, Italy, baseline characteristics are given in table 1 [10, 11]. In this cohort no data on lung function or COPD diagnosis was available.

Baseline characteristics of subjects included in the discovery cohort are presented in table 1. Males and females had a similar age and smoking status across the three academic sites. In the Groningen cohort, females had more severe airflow obstruction than males, *i.e.* lower FEV1 % predicted and FEV1/forced vital capacity (FVC) values, despite having smoked fewer pack-years. In the LAVAL and UBC cohorts, FEV1 % predicted and FEV1/FVC values were similar for males and females, but again with a lower number of pack-years smoked in females than males.

In the discovery cohort, a total of 51 159 probes were present on the array and 1225 probes, corresponding to 722 known genes, were differentially expressed between males and females (FDR<0.05), 609 being downregulated and 616 upregulated in females compared with males, with fold changes ranging from −166 to +207. Of the 609 downregulated probes in females, 28 were X-chromosome- and 74 Y-chromosome-linked, whereas 285 of the 616 upregulated probes in females were X-chromosome- and none Y-chromosome-linked. The expression of 392 of the 722 sex-associated genes described above were measured by the microarray used in the replication cohort (Illumina HumanHT-12 v4; Illumina, San Diego, CA, USA). A total of 199 (51%)

TABLE 1 Clinical characteristics of included subjects in the discovery and replication cohorts

| | Discovery cohort | | | | | | Replication cohorts | | |
| | Groningen | | Laval | | UBC | | Dugo et al. [11] | Morrow et al. [13] | |
	Male	Female	Male	Female	Male	Female		COPD	No COPD
Subjects	76	70	159	123	92	61	284	111	40
Age years	62.5 (55.0–70.3)	56.0 (51.0–61.0)**	66.0 (59.5–71.0)	60.0 (54.5–67.0)**	64.0 (59.0–70.3)	59.0 (53.0–71.0)	65[¶]	63[¶]	66[¶]
Males/females							202/82	52/59	15/25
COPD, yes/no	56/20	45/25	83/76	64/59	43/49	20/41			
Ex-/current smoking[#]	57/19	50/20	130/29	92/31	53/39	37/24	0/283/1	111/0/0	40/0/0
Pack-years	40.0 (29.0–50.0)	30.0 (20.8–35.6)*	48.0 (36.3–60.0)	40.0 (26.5–49.8)**	43.0 (34.5–69.5)	36.0 (21.8–47.3)**		61	34
FEV1 % pred	62.5 (39.0–78.3)	23.5 (18.5–79.3)*	78.4 (65.9–94.5)	79.3 (67.6–91.4)	79.2 (67.6–103.6)	87.5 (74.0–98.6)		27	99
FEV1/FVC %	54.5 (41.9–67.9)	43.3 (27.3–72.5)	67.3 (56.6–74.9)	68.9 (61.3–74.5)	68.3 (57.7–76.8)	73.2 (61.0–76.9)		30	80
T_{LCO} % pred			83.0 (68.5–93.5)	91.0 (72.3–111.8)	82.3 (67.8–91.6)	71.8 (57.5–85.3)			

Data are presented as n or median (interquartile range), unless stated otherwise. COPD: chronic obstructive pulmonary disease; FEV1: forced expiratory volume in 1 s; FVC: forced vital capacity; T_{LCO}: transfer factor of the lung for carbon monoxide. [#]: for the replication cohorts, smoking data are presented as ex-/current/never smoking; [¶]: mean. *: p<0.05; or **: p<0.01 for the difference in females compared with males within each discovery cohort (Mann–Whitney U-test).

out of these 392 genes could be replicated, *i.e.* were significantly associated with sex (nominal p-value<0.05) with a similar direction of effect. This is much more (>10 fold) than would be expected based on chance alone, which shows the robustness of our data. For genes upregulated in females, Gene Set Enrichment Analysis (GSEA, version 2.2.2, http://software.broadinstitute.org/gsea/index.jsp; using the Kyoto Encyclopedia of Genes and Genomes, KEGG, www.genome.jp/kegg/) demonstrated enrichment of genes involved in multiple biologically relevant pathways (FDR<0.1), the most significant being "translational initiation", "metabolism of xenobiotics by cytochrome P450" including the genes cytochrome P450 (*CYP*)*1A2*, *CYP1A1* and *CYP1B1* (figure 1a), "translational regulation", and "ribonucleoprotein complex subunit organization". For genes upregulated in males, significant pathways included "multicellular organismal macromolecule metabolic processes", "purinergic receptor signalling", and "negative regulation of calcium mediated signalling". The observation that females exhibit higher expression of genes involved in cytochrome P450 regulation, particularly *CYP1A2*, *CYP1A1* and *CYP1B1* (figure 1a), in lung tissue is of particular interest as it may render them more susceptible to the toxic effects of cigarette smoke and other environmental exposures (*e.g.* workplace or biomass). Indeed, it has been shown that CYP450 enzymes may convert cigarette smoke constituents like naphthalene to more toxic intermediate metabolites such as naphthalene-oxide [12].

To study how the COPD-related lung tissue gene expression changes differ between males and females, we performed a COPD*sex interaction analysis in the discovery cohort. This analysis showed a significant positive interaction between COPD and sex for 243 probes and a significant negative interaction for 200 probes, again at a nominal p-value <0.005. A positive interaction can be due to an increased gene expression in male COPD patients and no change in females with COPD, or a decrease in females with COPD without a change in males; a negative interaction can be due to a decrease in males with COPD and no change in females with COPD or an increase in females with COPD without a change in males. We replicated these findings in a publicly available dataset with lung tissue gene expression data available in 111 COPD patients and 40 control smokers (GSE76925; table 1) [13]. GSEA analysis showed statistically significant enrichment for genes with a positive interaction between COPD and sex in both cohorts (GSEA FDR<0.001).

The five most significant positively interacting genes were *Sphingomyelin synthase (SGMS)1, leucine rich repeat and Ig domain containing 2 (LINGO2), solute carrier family 25 member 1 (SLC25A1),*
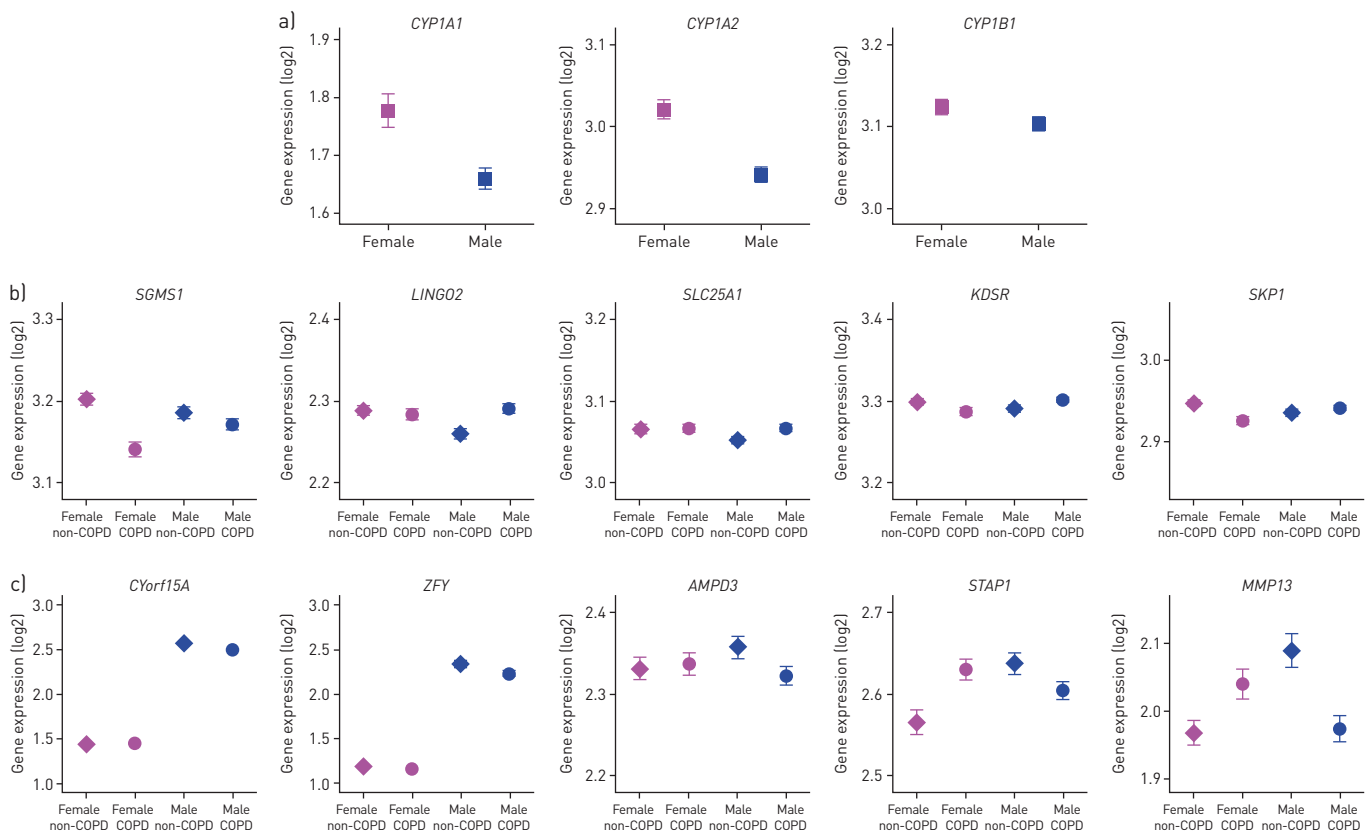

FIGURE 1 a) Differential expression of *CYP1A2*, *CYP1A1* and *CYP1B1* between males and females, and the five most significantly b) positively and c) negatively interacting genes. For example, the expression of *SGMS1* was increased in females *versus* without chronic obstructive pulmonary disease (COPD) without a change in males *versus* without COPD. Error bars indicate the standard error of the mean.

3-ketodihydrosphingosine reductase (KDSR) and *S-phase kinase-associated protein 1 (SKP1)* (figure 1b). The five most significant negatively interacting genes were *chromosome Y open reading frame 15A (CYorf15A), zinc finger protein Y-linked (ZFY), adenosine monophosphate deaminase 3 (AMPD3), signal transducing adaptor family member 1 (STAP1)* and *matrix metallopeptidase 13 (MMP13)* (figure 1c). GSEA analyses on the positively interacting genes revealed significant enrichment for pathways related to chemokine signalling and lymphocyte chemotaxis. The negatively interacting genes were enriched for KEGG pathways related to membrane assembly and biogenesis and sphingolipid biosynthesis. The most significantly positively interacting gene, downregulated with COPD in females, was *SGMS1,* a key regulator of sphingolipid metabolism. This finding is of special interest given our previous observation that lipids from the sphingolipid pathway are upregulated in COPD [14]. The initial step of sphingolipid metabolism is formation of ceramide in the endoplasmic reticulum, followed by transport to the Golgi complex *via* the ceramide transfer protein (CERT) and conversion to sphingomyelin by SGMS1 [15]. *Cert* mutant mice exhibit embryonic lethality due to mitochondrial degeneration [15]. Moreover, *sgms1* knockout mice suffer from mitochondrial dysfunction with higher oxidative stress [16]. The observation of decreased *SGMS1* expression in the lung tissue of females with COPD suggests that mitochondrial dysfunction and subsequent dysregulation of sphingolipid metabolism may be involved. Obviously, the other genes are also of interest to be explored further in future research.

Next, we analysed if the gene expression profile in lung tissue was differentially associated with COPD in males and females. In stratified analyses in either males or females, no genes were significantly downregulated or upregulated in association with COPD at the FDR cut-off level of <0.05. Therefore, to be able to compare differences, we used a more lenient cut-off of a nominal p-value <0.005. This resulted in a total of 137 probes being downregulated and 180 upregulated in males (fold changes ranging from −1.40 to +1.72), whereas 371 probes were downregualted and 464 upregulated in association with COPD in females (fold changes ranging from −1.55 to +1.85). Notably, there was very little overlap between males and females with respect to COPD-associated genes. Only five genes were downregulated (*FRMD4B, PRMT8, RAB35, SCD5,* and *VEGFA*) and two genes were upregulated (*FBLN5 and WHSC1L1*) with COPD in both males and females, compatible with earlier findings [9].

Our results provide further support for the notion that sex is an important factor that should be taken into account in future studies and clinical management of COPD. This adds to and extends previous findings that the bronchodilator response is larger in females than males with COPD, which may be due to the greater muscarinic M3 relative to M2 receptor expression in female compared with male lungs [17, 18].

Whereas pathway analyses did reveal statistically significant and biologically relevant sex-specific COPD-associated changes, there were no individual genes significant at the strict FDR cut-off of <0.05. So, future studies should be extended to increase power. Furthermore, it was not possible to replicate the COPD-associated sex differences in the replication cohort, since no data on lung function were available. Our study has some limitations as it was a *post hoc* approach and lung samples were not specifically collected for gene expression. In addition, baseline characteristics were not exactly similar between males and females, and a potential bias introduced by inconsistent concomitant medications across the groups cannot be ruled out. Finally, gene expression changes are not necessarily representative of protein expression and the latter was not investigated in our study.

In conclusion, we show that there is differential gene expression in the lungs of males and females of CYP450 enzymes that may be involved in the metabolism of toxicants present in smoke and/or air pollution. In addition, we demonstrate differential COPD-specific gene expression in males and females, with *SGMS1*, a gene involved in sphingolipid metabolism as a specific example suggesting that this metabolism may be dysregulated in females but not in males with COPD [17].

Maarten van den Berge[1,2,12], **Corry-Anke Brandsma**[2,3,12], **Alen Faiz**[1,2,4], **Maaike de Vries** [2,3], **Senani N. H. Rathnayake** [4], **Peter D. Paré**[5,6], **Don D. Sin**[5,6], **Yohan Bossé** [7,8], **Michel Laviolette**[7], **David C. Nickle**[9], **Ke Hao**[9], **Ma'en Obeidat**[5], **Tommaso A. Dragani** [10], **Francesca Colombo**[11], **Wim Timens**[2,3,13] and **Dirkje S. Postma**[1,2,13]

[1]University of Groningen, University Medical Center Groningen, Dept of Pulmonary Diseases, Groningen, The Netherlands. [2]University of Groningen, University Medical Center Groningen, Groningen Research Institute for Asthma and COPD, Groningen, The Netherlands. [3]University of Groningen, University Medical Center Groningen, Dept of Pathology and Medical Biology, Groningen, The Netherlands. [4]University of Technology Sydney, Respiratory Bioinformatics and Molecular Biology (RBMB), School of Life Sciences, Sydney, Australia. [5]The University of British Columbia, Center for Heart Lung Innovation, St Paul's Hospital, Vancouver, BC, Canada. [6]Respiratory Division, University of British Columbia, Vancouver, BC, Canada. [7]Institut universitaire de cardiologie et de pneumologie de Québec, Québec, QC, Canada. [8]Dept of Molecular Medicine, Laval University, Québec, QC, Canada. [9]Merck Research Laboratories, Boston, MA, USA. [10]Research Unit "Genetic Epidemiology and Pharmacogenomics", Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy. [11]Dept of Predictive and Preventive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy. [12] Shared first authorship; both authors contributed equally. [13]Shared last authorship; both authors contributed equally.

Correspondence: Maarten van den Berge, Dept of Pulmonary Diseases, University Medical Center Groningen, Hanzeplein 1, 9700 RB, Groningen, the Netherlands. E-mail: m.van.den.berge@umcg.nl

References

1   Watson L, Vestbo J, Postma DS, et al. Gender differences in the management and experience of Chronic Obstructive Pulmonary Disease. Respir Med 2004; 98: 1207–1213.
2   Silverman EK, Weiss ST, Drazen JM, et al. Gender-related differences in severe, early-onset chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2000; 162: 2152–2158.
3   Downs SH, Brandli O, Zellweger JP, et al. Accelerated decline in lung function in smoking women with airway obstruction: SAPALDIA 2 cohort study. Respir Res 2005; 6: 45.
4   Watson L, Schouten JP, Lofdahl CG, et al. Predictors of COPD symptoms: does the sex of the patient matter? Eur Respir J 2006; 28: 311–318.
5   Martinez FJ, Curtis JL, Sciurba F, et al. Sex differences in severe pulmonary emphysema. Am J Respir Crit Care Med 2007; 176: 243–252.
6   Watson L, Vonk JM, Lofdahl CG, et al. Predictors of lung function and its decline in mild to moderate COPD in association with gender: results from the Euroscop study. Respir Med 2006; 100: 746–753.
7   Camp PG, O'Donnell DE, Postma DS. Chronic obstructive pulmonary disease in men and women: myths and reality. Proc Am Thorac Soc 2009; 6: 535–538.
8   Hao K, Bosse Y, Nickle DC, et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. PLoS Genet 2012; 8: e1003029.
9   Brandsma C-A, van den Berge M, Postma DS, et al. A large lung gene expression study identifying fibulin-5 as a novel player in tissue repair in COPD. Thorax 2015; 70: 21–32.
10   Galvan A, Frullanti E, Anderlini M, et al. Gene expression signature of non-involved lung tissue associated with survival in lung adenocarcinoma patients. Carcinogenesis 2013; 34: 2767–2773.
11   Dugo M, Cotroneo CE, Lavoie-Charland E, et al. human lung tissue transcriptome: influence of sex and age. PLoS One 2016; 11: e0167460.
12   Chichester CH, Buckpitt AR, Chang A, et al. Metabolism and cytotoxicity of naphthalene and its metabolites in isolated murine Clara cells. Mol Pharmacol 1994; 45: 664–672.
13   Morrow JD, Zhou X, Lao T, et al. Functional interactors of three genome-wide association study genes are differentially expressed in severe chronic obstructive pulmonary disease lung tissue. Sci Rep 2017; 7: 44232.
14   Telenga ED, Hoffmann RF, t'Kindt R, et al. Untargeted lipidomic analysis in chronic obstructive pulmonary disease uncovering sphingolipids. Am J Respir Crit Care Med 2014; 190: 155–164.
15   Wang X, Rao RP, Kosakowska-Cholody T, et al. Mitochondrial degeneration and not apoptosis is the primary cause of embryonic lethality in ceramide transfer protein mutant mice. J Cell Biol 2009; 184: 143–158.
16   Yano M, Watanabe K, Yamamoto T, et al. Mitochondrial dysfunction and increased reactive oxygen species impair insulin secretion in sphingomyelin synthase 1-null mice. J Biol Chem 2011; 286: 3992–4002.
17   Li X, Obeidat M, Zhou G, et al. responsiveness to ipratropium bromide in male and female patients with mild to moderate chronic obstructive pulmonary disease. EBioMedicine 2017; 19: 139–145.
18   Calzetta L, Puxeddu E, Rogliani P. Gender-related responsiveness to the pharmacological treatment of COPD: a first step towards the personalized medicine. EBioMedicine 2017; 19: 14–15.