SUPPLEMENTARY MATERIAL

Title: **Hypermethylation of anti-oncogenic microRNA 7 is increased in emphysema patients**

Authors: Rocío Rosas-Alonso, Raúl Galera, Joan José Sánchez-Pascuala, Raquel Casitas, Miranda Burdiel, Elisabet Martínez-Cerón, Olga Vera, Carlos Rodriguez-Antolin, Olga Pernía, Javier De Castro, Francisco García-Rio, Inmaculada Ibanez de Caceres

METHODS

Study Subjects

Consecutive sampling was used to recruit 136 COPD patients from an academic medical center (La Paz University Hospital, Madrid, Spain). Inclusion criteria were: postbronchodilator FEV_1/FVC <0.7 and <lower limit of normal; history of smoking >20 pack-years; stable condition at inclusion, with no infection or exacerbation for at least 2 months; and optimal medical therapy for at least 8 weeks with no change. As control group, 30 smoker subjects (>20 pack-years) without evidence of airflow limitation (FEV_1/FVC > 0.7 and > lower limit of normal) were consecutively recruited from the lung cancer screening program of our center.

Exclusion criteria were a previous diagnosis of lung cancer or cancer of any other origin; history of asthma, bronchiectasis, tuberculosis, obliterative bronchiolitis or diffuse panbronchiolitis; previous use (3 months) of systemic corticosteroids; evidence of mouth lesions or any oropharyngeal disorder; clinical signs of acute heart failure; known unstable heart disease or disabling cognitive problems. The study was approved by the institutional ethics committee and each participant gave written informed consent (HULP-PI 2109).

Procedures

Clinical and Functional Evaluation

Anthropometric characteristics, smoking habit, comorbidities and current treatment were recorded. In COPD patients, dyspnea was graded using the modified Medical Research Council (mMRC) score,[1] and exacerbations requiring treatment with antibiotics, oral corticosteroids and/or hospitalization in the year prior to the study were also recorded.

Arterial blood gas values breathing room air were measured (Rapidpoint 405, Bayer, Munich, Germany). Spirometry was performed by means of a pneumotachograph, and static lung volumes were measured with a constant-volume body plethysmograph (MasterLab Body, Erich Jaeger GmbH, Würzburg, Germany), while diffusing capacity for carbon monoxide was

determined by the single-breath method (MasterLab Body) according to current recommendations.[2-4] Global Lung Initiative (GLI) predicted values were used for spirometry[5] and those of the European Coal and Steel Community[6] for lung volumes and diffusing capacity. Airflow limitation severity and COPD risk groups were stratified according to the GOLD statement.[7]

COPD phenotype classification

Clinical COPD phenotypes were classified according to the GesEPOC criteria.[8] The frequent exacerbator phenotype was defined by the presence of at least two exacerbations requiring antibiotics, oral steroids or one hospitalization in the preceding 12 months and separated by at least four weeks. Patients with <2 exacerbations in the previous year and cough and expectoration for 3 months of the year over two consecutive years were classified as chronic bronchitis phenotype. Patients with <2 exacerbations in the previous year without chronic bronchitis and with functional (reduced diffusing capacity) and/or radiological (computed tomography scan) evidence of emphysema were classified as emphysema phenotype. Asthma COPD overlap (ACO) was defined by the presence of one major criterion (previous history of asthma or a bronchodilator response higher than 15% or 400 ml) or two minor criteria (history of atopy, two separate bronchodilator responses higher than 12% and 200 ml, blood eosinophils >5%, or total IgE>100 IU).[9]

Analysis of miR-7 DNA methylation

Buccal epithelial swabs were obtained at the clinical consultation. The DNA from the 161 samples was eluted with PBS and digested by proteinase K (Invitrogen, Carlsbad, CA), followed by extraction with phenol/chloroform and used for methylation analysis of the specific CpG positions after bisulfite modification of DNA as previously reported.[10,11] The sodium bisulfited samples were used for quantitative methylation-specific polymerase chain reaction (qMSP) by

designing two probes (FAM and VIC), to specifically recognize methylated or unmethylated CpG positions, making it possible to measure the percentage of methylated DNA in a unique PCR reaction. We used the following primer/probe set for methylated reaction: **F**: 5'-GGGTGGGGTTTTTTAAGAATC-3'; **R**: 5'-ACATTCTCCTCCTTCGATCG-3'; Probe: 5'-**FAM**-ACCCCTCTTCGTTCTCGAT-3'; and for unmethylation: **F**: 5'-GGGGTGGGGTTTTTTAAGAATT-3'; **R**: 5'-ATAACATTCTCCTCCTTCAATCA-3'; Probe: 5'-**VIC**-ACCCCTCTTCATTCTCAAT-3', using the following settings: 50°C-2 min; 95°C -15 min; and 40 cycles of 94°C- 50 seg, 60°C-1 min. All assays were performed in duplicate using the QuantiTect Multiplex PCR Kit (Qiagen, USA) and the HT7900 Applied Biosystems. Each set of modified DNA included DNA from healthy donors as a negative control and DNA methylated *in vitro* with Sss I methylase (New England Biolabs, Beverly, MA) to be used as a positive control. The percentage of methylation of each sample was calculated as previously reported.[11]

Statistical analysis

Data are summarized as mean ± SD for continuous variables, while frequencies (percentages) are used for categorical variables. The normality of the distribution of variables was tested using the Kolmogorov-Smirnov and Shapiro-Wilk tests. Differences between study groups were analyzed using analysis of variance with post-hoc comparisons by the Bonferroni test or by the chi-square test. The effect of the possible confounding factors was assessed using generalized linear models.[12] We constructed a multivariate model, including COPD phenotype as fixed factor and gender, age, BMI, smoking status, airflow limitation severity, GOLD risk group, and treatment with inhaled corticosteroids as covariates. The link function used was the identity. We chose normal distribution because it was more fitting than inverse Gaussian or gamma distribution, according to the plausibility criteria, Pearson's chi-square and analysis of deviance.

The relationships between methylated miR-7 levels and clinical/functional characteristics of COPD patients were determined using Pearson's linear bivariate correlation. Gender, age, smoking habit, airflow limitation severity and significantly related variables were then used in a multiple linear regression analysis to identify independent determinants of methylated miR-7 levels. In this analysis, the stepwise method was used to include or remove one independent variable at each step, based on the probability of F (entry: 0.05; removal: 0.10). Other aspects explored included residual standard deviation, changes in the distribution of the residuals and homogeneity of the variance over the predictors. The assumptions of linearity and distributional normality were confirmed for all variables. A histogram of residuals and a normal probability plot of residuals were used to test for normality. Homoscedasticity was explored by scatter plots of the standardized residuals on the standardized predicted values and by Levene's test for equality of variances. When heteroscedasticity was suspected, regression coefficients were computed using weighted least squares. In all regression models, multicollinearity was verified by the determination of the variance inflation factor <3 for each predictor.

All analyses were performed using SPSS 13.0 software. A $p<0.05$ was accepted as the minimum level of statistical significance.

REFERENCES

1. Bestall JC, Paul EA, Garrod R, Garnham R, Jones PW, Wedzicha JA. Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax.* 1999; **54**: 581-6.
2. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardization of spirometry. *Eur. Respir. J.* 2005; **26**:319-38.
3. Wanger J, Clausen JL, Coates A, Pedersen OF, Brusasco V, Burgos F, et al. Standardization of the measurement of lung volumes. *Eur. Respir. J.* 2005; **26**: 511-22.
4. Macintyre N, Crapo RO, Viegi G, Johnson DC, van der Grinten CP, Brusasco V, et al. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. *Eur. Respir. J.* 2005; **26**: 720-35.
5. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur. Respir. J.* 2012; **40**: 1324-43.
6. Quanjer PH, Trammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests. European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur. Respir. J.* 1993; **6**(Suppl. 16): 5s-40s.
7. Vogelmeier CF, Criner GJ, Martínez FJ, Anzueto A, Barnes PJ, Bourbeau J, et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report: GOLD Executive Summary. *Arch. Bronconeumol.* 2017; **53**: 128-49.
8. Miravitlles M, Soler-Cataluña JJ, Calle M, Molina J, Almagro P, Quintano JA, et al. Spanish Guidelines for Management of Chronic Obstructive Pulmonary Disease (GesEPOC) 2017. Pharmacological Treatment of Stable Phase. *Arch. Bronconeumol.* 2017; **53**: 324-35.

9. Cosio BG, Soriano JB, López-Campos JL, Calle-Rubio M, Soler-Cataluña JJ, de-Torres JP, et al. Defining the asthma-COPD overlap syndrome in a COPD cohort. *Chest.* 2016; **149**: 45-52.

10. Vera O, Jimenez J, Pernia O, Rodriguez-Antolin C, Rodriguez C, Sanchez Cabo F, et al. DNA methylation of miR-7 is a mechanism involved in platinum response through MAFG overexpression in cancer cells. *Theranostics.* 2017; **7**: 4118-34.

11. Vera-Puente O, Rodriguez-Antolin C, Salgado-Figueroa A, Michalska P, Perna O, Reid BM, et al. MAFG is a potential therapeutic target to restore chemosensitivity in cisplatin-resistant cancer cells by increasing reactive oxygen species. *Transl. Res.* 2018; **200**: 1-17.

12. Snijders T, Bosker R. *Mutilevel analysis: an introduction to basic and advanced multilevel modeling*. Sage, London, 1999.

Table S1. Univariate linear regression models of predictors of methylated miR-7 in COPD patients*

	r	95% CI	p
Age, yr	0.224	0.048 to 0.386	0.013
BMI, Kg/m^2	0.064	-0.117 to 0.240	0.484
Pack-years	-0.124	-0.321 to 0.083	0.241
FVC, % pred.	0.074	-0.106 to 0.249	0.423
FEV_1, % pred.	0.281	0.108 to 0.438	0.002
FEV_1/FVC	0.255	0.079 to 0.415	0.005
TLC, % pred.	-0.089	-0.364 to 0.200	0.546
FRC, % pred.	0.376	0.103 to 0.596	0.009
FRC/TLC, %	0.532	0.292 to 0.709	<0.001
RV, % pred.	0.228	-0.060 to 0.481	0.120
RV/TLC, %	0.142	-0.201 to 0.454	0.414
DLCO, % pred.	-0.450	-0.637 to -0.212	<0.001
PaO_2, mmHg	0.038	-0.208 to 0.279	0.767
C reactive protein, ng/dl	-0.067	-0.314 to 0.188	0.609

Definition of abbreviations: BMI=body mass index; FVC=forced vital capacity; FEV_1=forced expiratory volume at 1 second; TLC=total lung capacity; FRC=functional residual capacity; RV=residual volume; DLCO=diffusion capacity for carbon monoxide; PaO2=arterial pressure of oxygen

*Data recorded are Pearson's correlation coefficient, 95% confidence interval of the correlation and p value.