Differentiating asthma from chronic obstructive pulmonary disease via their metabolomic signatures – Preliminary work

Ricardo Gomes [1], Guillaume L. Erny [2], Arminda Alves [2], Nuno Neuparth [3,4], Pedro Carreiro-Martins [3,4], João Gaspar Marques [3,4], Patrícia Gomes-Alves [1]*

[1]UniMS – Mass Spectrometry Unit, ITQB/IBET, Oeiras, Portugal; [2] LEPABE - Laboratory for Process Engineering, Environment, Biotechnology and Energy, Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal; [3] CEDOC - Integrated Pathophysiological Mechanisms Research Group. NOVA Medical School / Faculdade de Ciências Médicas, Rua do Instituto Bacteriológico n.º 5, 1150-190 Lisboa, Portugal; [4] Serviço de Imunoalergologia, Hospital de Dona Estefânia, Centro Hospitalar de Lisboa Central, EPE, Rua Jacinta Marto, Lisboa, Portugal;

* ragomes@ibet.pt; palves@ibet.pt

Introduction

Chronic obstructive pulmonary disease (COPD) diagnosis and subsequent management is largely based on clinical assessment which may lead to misdiagnosis that occurs, in particular, in primary care settings.[1] Clinical assessment does not always accurately differentiate COPD from asthma or from asthma-COPD overlap leading inadequate treatment and increased risk of morbidity and mortality.

Exhaled breath condensates (EBCs) are obtained by condensation of gases and droplets released during exhalation. Traditionally, pulmonologists focused on respiratory gases in these plumes, but it has been recently demonstrated the EBCs are rich sources of biomarkers and can reflect inflammation processes that are occurring within the lungs. However, biomarkers are often found a very low concentration.

This study aims to establish the metabolomics signature of COPD and asthma patients using liquid chromatography hyphenated to high-resolution mass spectrometry. Such untargeted analysis will rely on a computer-assisted MS analysis tool (CAMSAT) called Finnee to mine for features of very low intensities (peaks) and measure key chromatographic parameters (intensities, migration time and accurate masses).[2]

Methods

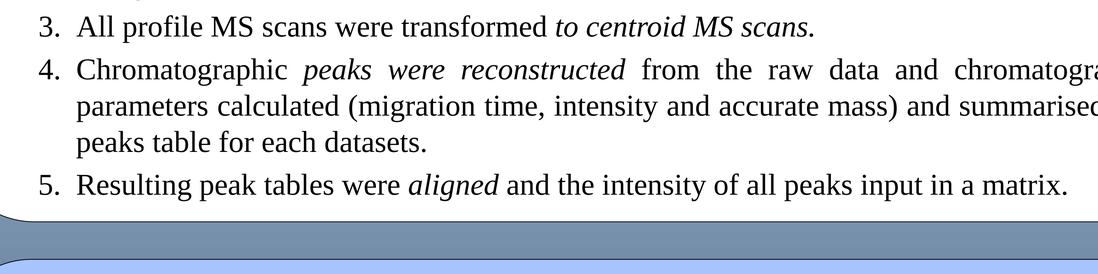

Figure 1: Schematic representation of the methodology developed for this work

The following methodologies were devised for this preliminary work:

- **EBC samples and clinical assessment**
  EBCs from 15 patients (5 healthy, 5 with asthma symptoms and 5 with COPD symptoms, as assessed by the Serviço de Imunoalergologia, Hospital Dona Estefânia) have been collected.

- **Separation parameters**
  Liquid chromatography: HALO C18, 2.7 µm, 90 Å, 0.5x50 mm column using a 30 min gradient from 5-95% B (solvent A: water + 0.1% FA; solvent B: Acetonitrile + 0.1% FA).
  Mass spectrometry: MS analysis was performed using a Sciex TripleTOF6600 with the DuoSpray Ion source. MS spectra accumulation time of 210 msec was used.

- **Data post-processing**
  Data were further analysed using *Finnee* (https://finneeblog.wordpress.com/) a Matlab toolbox that aims in mining chromatographic peaks in LC-HRMS datasets in an untargeted approach. The following steps were used:
  1. The original files were *converted to the mzML* format using msConvert.
  2. mzML files were then transformed to Matlab objects, corrected for *baseline drift and the background noise* removed.[3]
  3. All profile MS scans were transformed *to centroid MS scans.*
  4. Chromatographic *peaks were reconstructed* from the raw data and chromatographic parameters calculated (migration time, intensity and accurate mass) and summarised in a peaks table for each datasets.
  5. Resulting peak tables were *aligned* and the intensity of all peaks input in a matrix.

Results

Description of steps	Running Time	Size
Original file (wiff file)		*NA*
Conversion to mzML	15 min	*NA*
Conversion to Matlab Object	25 min	2.7 GB
Baseline correction	35 min	1.0 GB
Noise removal	15 min	8 MB
Conversion of profile MS scans to centroid scans	0.5 min	800 KB
Mining for chromatographic features and characterization	5 sec	*NA*

See Figure 2

See Figure 3

Table 1: Running time for the different computerised step for a single dataset (CPU: Intel Core i7 7700HQ @ 2.80GHz with 16.0GB of RAM)

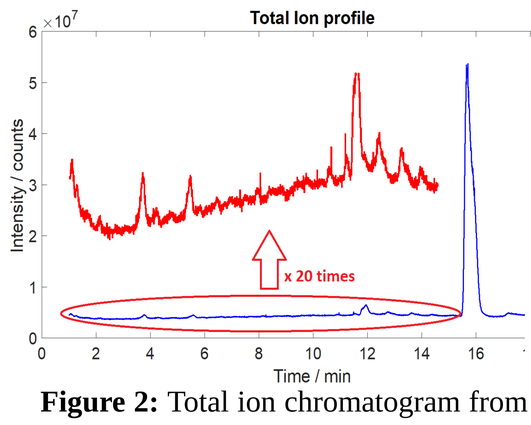

Figure 2: Total ion chromatogram from an asthmatic EBC sample

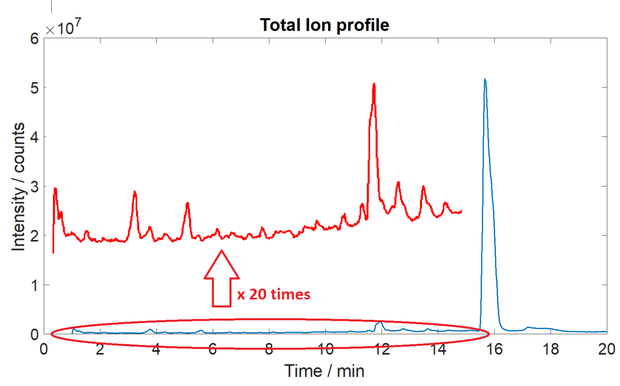

Figure 3: Total ion chromatogram after correcting for baseline drift and background noise

Main achievements:

- ✓ **1200 to 1500 peaks per dataset**
- ✓ **Intensities ranging from 500 to 5,000,000**
- ✓ **After peaks matching, 1173 selected peaks (4 to 15 occurrences in all datasets)**

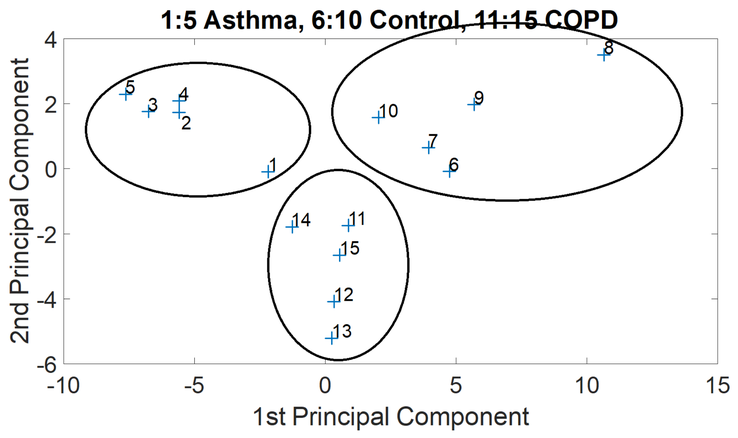

Figure 4: Variance weighted principal component analysis (PCA) using an intensity matrix with 15 observations (5 controls, 5 asthma and 5 COPD) and 1173 variables (peak intensities). Intensities were normalised before the PCA by the sum of intensities for each observation.

References

[1] Lusuardi M, De Benedetto F, Paggiaro P, et al. A randomized controlled trial on office spirometry in asthma and COPD in standard general practice: data from spirometry in Asthma and COPD: a comparative evaluation Italian study. *Chest* 2006; 129:844-52.

[2] Erny, GL, Acunha, T, Simó, C, Cifuentes, A, Alves, A. Finnee—A Matlab toolbox for separation techniques hyphenated high resolution mass spectrometry dataset. *Chem. Int. Lab.*; 155: 138-144.

[3] Erny, GL, Acunha, T, Simó, C, Cifuentes, A, Alves, A. Background correction in separation techniques hyphenated to high-resolution mass spectrometry–Thorough correction with mass spectrometry scans recorded as profile spectra. *J. Chrom. A;* 1492: 98-105

Acknowledgements

This work was financially supported by the projects: (i) POCI-01-0145-FEDER-006939 (Laboratory for Process Engineering, Environment, Biotechnology and Energy – UID/EQU/00511/2013) funded by the European Regional Development Fund (ERDF), through COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI) and with financial support from FCT/MCTES through national funds (PIDDAC). (ii) NORTE-01-0145-FEDER-000005 – LEPABE-2-ECO-INNOVATION, supported by North Portugal Regional Operational Programme (NORTE 2020), under the Portugal 2020 Partnership Agreement, through the European Regional Development Fund (ERDF).(iii) FEDER funds through the Operational Program for Human Potential and by National Funds through FCT under the project IF/00528/2013. (iv) AstraZeneca – Projecto OLDER (CEDOC/2015/59); (v) iNOVA4Health - UID/Multi/04462/2013, financially supported by FCT/Ministério da Educação e Ciência, and co-funded by FEDER under the PT2020 Partnership Agreement.

Conclusions

- ✓ **Finnee allows extracting peak of very low intensity, ~1200 peaks whose intensities range four order of magnitudes were recover from each dataset;**
- ✓ **After PCA, asthma, COPD and control EBCs samples were accurately differentiated, without any samples preparation techniques;**
- ✓ **Simple and inexpensive sample pre-treatment techniques should be investigated to improve the quantity and quality of information;**
- ✓ **Specific biomarkers should now be identified and fully characterised;**
- ✓ **The study should be extended to include various asthma phenotypes and endotypes.**