A multiplatform metabolic phenotyping approach integrated with pathway mapping to identify biochemical differences between healthy smokers and non-smokers

Manuja Kaluarachchi1, Claire Boullange2, Isabel Garcia Perez2, John Lindon3, Emmanuel Minet1

1 British American Tobacco R&D CentreSouthampton, S015 8TL UK.
2 Metabometrix Ltd, Bio-incubator, Prince Consort Road, South Kensington, London SW7 2BP UK
3 Correspondence: emmanuel_minet@bat.com

Introduction
The development of next generation tobacco and nicotine products offers a significant opportunity to reduce the burden of tobacco use on population health, but epidemiological data are currently lacking. Systems toxicology consists of the integrative study of the biological study of the toxicities caused by toxicants using untargeted in vitro and in vivo assays combined with computational tools to determine possible toxicological outcomes. In this context, systems toxicology could form part of a weight of evidence approach for novel tobacco product risk assessment.

Objective
Our aim is to create a comprehensive map of tobacco smoking-related metabolable perturbations that can be interrogated for comparison against novel potentially reduced risk tobacco and nicotine products.

Materials and Methods
In this study we performed a multiplatform comparison (NMR, UPLC-MS RI-, HIILC-MS) of the serum metabolome/lipidome of smokers (n=66) and never smokers (n=57) with profiling of 80 serum mRNAs and biomarkers of exposure in urine. Multivariate and univariate statistical modeling allowing for confounder effects was performed.

Results
Representative NMR, RP-UPLC-MS, and HiILC-UPLC-MS spectral profile of smokers

Figure 1: Metabolic phenotyping data of blood serum samples from smokers. Median 1H-NMR (A) Noisy and (B) Cqmp spectral acquisition. Base peak intensity (BPI) chromatograms from (C) positive electrospray ionization mode (ESI+) and (D) negative electrospray ionization mode (ESI−) from the lipidomics-UPLC-MS analysis. (E) ESI+ and (F) ESI− from the HiILC-UPLC-MS analysis. Key PC: phosphatidylcholines, PG: phosphatidylglycerols, PE: phosphatidylethanolamines, SM: sphingomyelins, DG: diacylglycerols, CE: cholesterol esters, TG: triacylglycerols, AA: amino acids.

Figure 2: PCA Scores plots of the metabolic data from 1H-NMR and UPLC-MS platforms. Each spot on the scores plot represents an individual.

Figure 3: Differences in the metabolomics data between smoker and never smoker groups after correction for confounders (BMI, gender, age and drug intake). (A) Manhattan plots of the NMR cpg data. Significant p-value threshold of pFDR<0.05 (Benjamini-Hochberg false discovery rate) is marked with the dotted line. (B) Differential significance plots for all four UPLC-MS data sets showing the retention time on the x-axis and the nC value on the y-axis for each significant pFDR<0.05 feature. The colour of the spot relates to the direction of the correlation (up in smokers vs never smokers: orange and down in smokers vs never smokers: blue). The shading shows the strength of the correlation (the darker the colour the more significant). (C) Histogram showing the 12 lipoprofile subclasses (out of 105 measured) being significantly different: pFDR<0.05 between never smokers (green) and smokers (red). (D) Volcano plot showing the difference in mRNA expression between smokers and never smokers is plotted on the x-axis (log2 scale) (dotted vertical line marks the 2-fold increase threshold), and (FDR)-Bonferroni adjusted significance pFDR<0.05 dotted horizontal line is plotted on the y-axis (log10 scale).

Table 1: Clinical study demographics.

<table>
<thead>
<tr>
<th></th>
<th>Smokers</th>
<th>Never Smokers</th>
<th>pFDR&lt;0.01</th>
<th>pFDR&lt;0.05</th>
<th>&amp; FC&gt;10%↘</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnicity</td>
<td>46 (64%)</td>
<td>57 (85%)</td>
<td>0.026</td>
<td>0.046</td>
<td>0.067</td>
</tr>
<tr>
<td>Gender</td>
<td>28 (40%)</td>
<td>20 (30%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>36.9 (1.0)</td>
<td>42.0 (2.00)</td>
<td>0.026</td>
<td>0.046</td>
<td>0.067</td>
</tr>
<tr>
<td>Median (min. max.)</td>
<td>41.8 (26.36)</td>
<td>62.0 (26.90)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cigarette Consumption</td>
<td>56 (84%)</td>
<td>52 (76%)</td>
<td>0.026</td>
<td>0.046</td>
<td>0.067</td>
</tr>
<tr>
<td>Nicotine use</td>
<td>2 (30%)</td>
<td>0 (0%)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4: Metaboscanal enrichment plot for metabolites significant at pFDR<0.05. Top perturbed pathways based on impact and log(p-value) are labelled on the graph. The -log(p-value) is the enrichment score. The Impact score (0 to 1) indicated the pathway topological importance of the metabolites.

Conclusions
- Applying a multiplatform metabolomics, lipidomics, and epigenetics approach produced a rich dataset, which in combination with a knowledge-base enrichment analysis gave mechanistic insights into adverse biological events potentially associated with tobacco-related diseases in healthy smokers.
- Pathway perturbations were identified associated with EGF and P3-NC signaling, which is involved in non-small cell lung cancer and cardiovascular diseases development.
- The enrichment analysis is in agreement with previously published single platform metabolomics, lipidomics, and biomarkers studies in smokers and the multi-platform approach gives more granularity than single platform approaches.
- The reversibility of these alterations in metabolic pathways should be investigated to use as risk assessment benchmarks in cessation studies.

Reference:

Conflict of interest statement: This study was a collaboration between BAT and Metabometrix Ltd. The study was funded by BAT at Metabometrix Ltd as a commercial contract.

Clinical Study compliance statement: The study protocol and informed consent forms were approved by the Ethics Committee of the Antwerp Heart Institute, Germany, and the clinical study was conducted in accordance with the World Medical Association Declaration of Helsinki (World Medical Association, 2018). The study was registered in the Current Controlled Trials database under the reference ISRCTN18328686.
A multiplatform metabolic phenotyping approach integrated with pathway mapping to identify biochemical differences between healthy smokers and non-smokers

Manuja R. Kaluarachchi, Claire Boulangé, Isabel Garcia-Perez, John C. Lindon and Emmanuel Minet

1 British American Tobacco R&D Centre, Southampton, SO15 8TL UK, 2 Metabometrix Ltd, Bio-incubator, South Kensington, London SW7 2BP UK

Correspondence: emmanuel_minet@bat.com

A biological system can be simplified to a network of genes, proteins, metabolites interacting with each other in equilibrium. A toxicological stress can lead to a temporary or permanent perturbation of this network. Thus if a detailed signature of biological perturbations caused by an exposure event can be obtained and mapped to biological pathways it becomes possible to model the causal relationship between exposure and adverse events leading to diseases. In the context of risk assessment, creating a comprehensive map of the perturbed biological functions caused by tobacco smoking could be useful for comparison against novel tobacco/nicotine products.

The objective of this study was to perform a comprehensive multiplatform comparison (NMR, UPLC-MS RP+/−, UPLC-MS HILIC+/−) of the serum metabolome/lipidome of smokers (n=56) and non-smokers (n=57).

Multivariate statistical modelling and univariate statistic allowing for confounder variables was performed. From this a panel of 110 metabolic biomarkers was produced that distinguish the two classes. 88 were tentatively identified of which 62 varied by 10% or more between the two groups at pFDR<0.01. Some of the strongest changes were observed in smokers for L-tryptophan (−1.6x), hypoxanthine (+2.2x), and 1-acylglycerophosphocholine (−1.7x).

The identified metabolites were used as input for metabolic pathway enrichment analysis using IPA Ingenuity® and the Reactome database to predict adverse biological events. Lipid peroxidation, glutathione metabolism, and inflammatory response were predicted. Those are known adverse biological events in smokers indicating that combining a metabolomics approach with knowledge-base mapping can give mechanistic insights into disease development. A similar approach can be applied to next generation tobacco/nicotine products for risk assessment.

Related publication:

Conflict of interest statement: This study was a collaboration between BAT and Metabometrix Ltd. The study was funded by BAT to Metabometrix Ltd. as a commercial contract.

Clinical Study compliance statement: The study protocol and informed consent forms were approved by the Ethics Committee of the Ärztekammer Hamburg, Germany and the clinical study was conducted in accordance with the World Medical Association Declaration of Helsinki (World Medical Association, 2004). The study was registered in the Current Controlled Trials database under the reference ISRCTN12306288.