Objective: Systems toxicology consists of the integrative study of the biological perturbations caused by toxins using untargeted in vitro and in vivo screens combined with computational tools to determine possible toxicological outcomes. 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. In this context, systems toxicology could form part of a weight of evidence approach for novel tobacco product risk assessment. Our objective is to create a comprehensive map of tobacco smoking-related metabolite perturbations that can be interrogated for comparison against novel potentially reduced-risk tobacco and nicotine products.

Method: In this study we performed a multiphasic comparison (NMR, UPLC/MS RP+/-, Hilic+/-) of the serum metabolome/lipidome of smokers (n=56) and never smokers (n=57) with profiling of 80 serum miRNA and biomarkers of exposure in urine. Multivariate and univariate statistical modeling allowing for confounder variables was performed.

Results & conclusions: 1. Representative NMR, RP-UPLC-MS, and Hilic-UPLC-MS spectral profile of smokers

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

2. Group separation based on the metabolic profile obtained for each analytical platform according to gender and smoking status

- **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. The data have been coloured by gender (female, pink and male, blue) or smoking status (smoker, red and never smoker, green).

3. Differential metabolomics, lipidomics, and miRNA analysis in the serum of smokers and never smokers

- **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 cpgm data. Significant p-value threshold of FDR (Benjamini-Hochberg false discovery rate)<0.05 is marked with the dotted line. (B) Differentially significant plots for all four UPLC-MS data sets showing the retention time on the x-axis and the m/z 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 lipid subclasses being significantly different (pFDR<0.05) between never smokers (green) and smokers (red). (D) Violin plot showing the difference in miRNA expression between smokers and never smokers is plotted on the x-axis (log10 scale) (dotted vertical line marks the 0.05 fold change threshold), and (FDR)-Bonferroni adjusted significance (pFDR<0.05 dotted horizontal line) is plotted on the p-axis (−log10 scale).

4. Enrichment analysis: A panel of 110 metabolites and one differentially expressed miRNA was produced that distinguish the smokers and never smokers. A total of 62 metabolites was tentatively identified of which 59 varied by 10% or more at pFDR<0.05. Hidrolive, glutathione, PS22.6/12.5 and PE20.4/17.1 were correlated with the urine total nicotine dose. The metabolites were used as input for metabolic pathway enrichment using KEGG (Figure 4A) and Metaanalyst (Figure 4B). Co-abundance of ceramides and GSH was observed in smokers which has been associated with COPD. Similarly, perturbations of the choline pathway, ceramides, miR-124, and central carbon metabolism were identified, which have been involved in EGF-R-RAS/PI3K-AKT oncogenic signalling.

- **Figure 4:** Enrichment analysis with KEGG (A) and metaanalyst (B)

5. Conclusion:  
- Applying a multiphasic metabolomics, lipidomics, epigenetics and biomarkers of exposure 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.
- Metabolic perturbations were identified associated with EGF-R and PG-AKT signalling which is involved in non-small cell lung cancer.
- Enrichment analysis is in agreement with previously published single platform metabolomics, lipidomics, and biomarkers studies in smokers.
- The reversibility of these alterations in metabolic pathways should be investigated to use as risk assessment benchmark in studies where smokers are switched to novel tobacco and nicotine delivery products.