

Metabolomics Analysis Report

## Experiment EX00257 –

## Life course exposures and diet

## Experiment EX00181 –

## ELEMENT preliminary data analysis

Client: Dr. Karen Peterson

Author: Dr. Alexander Raskind

August 6, 2014

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 5458 Brehm Tower,1000 Wall Street, Ann Arbor, MI 48105-5714 U.S.A. |  | p. (734)232-0846f. (734) 232-8175 |  | araskind@umich.eduhttp://mrc2.umich.edu/ |

Table of Contents

1. [Introduction 2](#_Toc394080742)
2. [Study Summary 2](#_Toc394080743)
3. [Results 4](#_Toc394080744)
4. [Discussion 8](#_Toc394080745)
5. [Materials and methods 9](#_Toc394080746)

# Introduction

## Untargeted analysis overview

Untargeted analysis is aimed at the broadest possible coverage of metabolome to allow interpretation and hypothesis generation not limited to the existing knowledge and pre-conceived ideas. The downside of this approach is that analysis conditions cannot be optimal for all the compounds in a complex mixture. For this reason, untargeted metabolomics results may not be considered quantitative and thus the only legitimate way of interpretation is through comparison between the groups of samples.

## What is included in the report

* Report document (this)
* Data tables in Excel format

# Study Summary

The summary is for “ELEMENT preliminary data analysis” experiment, the details of the “Life course exposures and diet” experiment were not provided to the core. Experimental design for both experiments is not available as well.

#### ELEMENT preliminary data analysis summary

Human serum specimens were collected via the P20 Children's Environmental Health Center (UM-SPH) research project (NIEHS/EPA P20 ES018171/RD834800). Participants in this project are recruited from the longitudinal birth cohort study, ELEMENT (Early Life Exposures in Mexico to ENvironmental Toxicants). Archived data on these samples include dietary data (FFQ), anthropometry and exposure measures from multiple time points ranging from the prenatal to early pubertal life stages. Epigenetic analyses have also been performed on cord and current venous blood samples for these 40 subjects, thereby providing rare, longitudinal analyses of the epigenome.

# Results

## Data report structure

The report document is accompanied by the three Excel files containing data obtained by the Core and additional material necessary for further analysis and interpretation of results. Two complementary data sets were generated from the same samples - one in positive Mass Spectrometry (MS) mode and another in negative. In positive mode molecules that form positively charged ions are analyzed and in negative mode – those that form negatively charged ions. Thus combination of two modes gives a better coverage of metabolome than any single mode. There is a certain overlap between compounds detected in positive and negative mode since many molecules may form both positive and negative ions depending on conditions. This overlap is often useful to verify the results by looking if the same trend for certain compound is observed in both modes. Positive and negative mode data were analyzed separately and results are collated in **EX00257-PositiveMode.xlsx** and **EX00257‑NegativeMode.xlsx** files respectively. The quality control (QC) data for both modes are in **EX00257-QC.xlsx**. Summary of quality control data is also given in this report.

The data files are organized in several worksheets, each worksheet name is prefixed by NEG- for negative mode and POS- for positive. The list of worksheets is as follows:

* Setup – contains map of experimental design to sample file names
* Annotations – list of detected compounds (identified and unknown) with available information
* Log2 – list of log2 transformed peak areas per compound per sample
* Raw – list of raw peak areas per compound per sample

#### Quantitative results format notes

The numbers are MS ion counts (raw or log2 transformed). They are proportional both to compound concentration and ability to be ionized and for this reason cannot be used to compare absolute concentrations of different compounds.

The only legitimate comparisons you can make using our data are between same compounds in different samples or between ratios of different compounds within each sample.

You have complete data set, which means that all detected compounds are reported. There is a certain cutoff for detection, but it is quite low and if compound is missing in one of the samples, while in others it is present in substantial amount, it is almost certainly an artifact of search procedure. If certain compounds are important for your interpretation, but you think that there are missing or questionable values in certain samples we will have to manually go through original raw data to investigate the underlying reasons.

## Compound identification

Report includes both identified and unidentified compounds (unknowns). Identification is based on comparing observed mass and retention time to the in-house library of ~750 compounds. The degree of confidence in this type of identification is high, but possibility of misidentification still exists, especially for structural isomers with very similar chromatographic behavior. Compounds not matched to the library are assigned putative molecular formula, where possible. Identification results are provided in the “annotation” worksheets of data files.

|  |  |  |
| --- | --- | --- |
| Internal standard | Positive mode RSD% | Negative mode RSD% |
| Creatinine-(guanidino-13C) | 6.6 | 12.4 |
| Epibrassinolide | 1.2 | 1.5 |
| Gibberellic acid | 13.3 | 0.5 |
| Jasmonic acid | 2.1 | 0.3 |
| L-[15N] Anthranilic acid | 0.3 | 8.8 |
| L-Arginine-(guanidineimino-15N2) | 8.6 | 34 |
| L-Tryptophan-15N2 | 4.4 | 6 |
| Palmitic acid-13,13,14,14,15,15,16,16,16-d9 | 11.7 | 21 |
| Thymine-d4 (methyl-d3,6-d1) | 9.3 | 26.4 |
| Zeatin | 1.7 | 0.3 |

## Quality control data

Quality of analysis is accessed by visual inspection of the chromatographic traces (total ion chromatograms) and relative quantification of the internal standards. Quality control data are summarized in Table 2. Comprehensive QC data are enclosed in **EX00257-QC.xlsx** file. Analysis reproducibility is also evaluated using comparison of multiple injections of the pooled sample created by combining aliquots of all the samples included in the set. Figures 01 and 02 show the distribution of relative standard deviation for all detected compounds for pooled samples only (blue) vs all samples in data set (pink). RSD distribution for pooled samples has a slightly sharper maximum around ~25% value in both modes while overall RSD maximum is ~40%. The relatively high RSD for pooled samples is due to large sample number which required analysis to be split in several batches.

**Table 2 Internal standard reproducibility**

##

Figure 0.3 Density distribution of peak area relative standard deviation for negative mode analysis.

Figure 0.2 Density distribution of peak area relative standard deviation for positive mode analysis.

Figure 0.1 Principal Component Analysis

# Discussion

Quality control data suggest that analysis method was sufficiently stable and reproducible across all samples. There is however a noticeable batch effect revealed by PCA analysis (not shown). The nature of batch effect is not completely clear. Sophisticated statistical analysis will be required to take batch effect into account for data interpretation. Another problem is that since experimental design data were not provided samples were split into batches in an arbitrary way although strictly speaking the proper block design would be necessary. Biochemical interpretation of the results is complicated since relatively few compounds are identified. We are currently working on expanding our compound library and may be able to provide a more complete list of identified compounds once the new library would be in place. We also may provide assistance in identification of unknowns if your analysis of the data will reveal such compounds to be of a particular interest.

# Materials and methods

## Sample preparation

**Extraction solvent:** MAA (Methanol : Acetonitrile : Acetone 1:1:1) with internal standards; for 200 samples 100 ml extraction solvent and 4 ml internal standard mixture.

**Reconstitution solvent:** Methanol : H2O 2 : 98

**Procedure:**

* Pipette 100 µL of medium from the sample vials into a labeled 1.5 ml micro-centrifuge tube
* Add 400 µL of extraction solvent to all tubes.
* Vortex for 5 min, then let sit 30 min at 4°C. Vortex again and let sit at -20°C for 1 h, then centrifuge 10 min at 15,000 rpm.
* Transfer 250 µL of supernatant to labeled labelled 1.5 ml micro-centrifuge tube.
* Dry all samples under nitrogen stream at room temperature.
* Reconstitute in 100 µL of reconstitution solvent.
* Vortex for 5 min, centrifuge 5 min at 15,000 rpm.
* Transfer supernatant to autosampler vials with inserts, discard pellets.

## LC-MS

### Chromatography

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1290 Infinity Binary LC System from Agilent is used for chromatographic separation together with Waters Acquity UPLC HSS T3 1.8 µm 2.1 x 100 mm column in connection with a Water Acquity UPLC HSS T3 1.8 µm VanGuard pre-column.* Data acquisition: time 27 min
* System equilibration time: 7 min
* Total run length: 34 min
* Flow rate: 0.45 ml/min
* Solvent A: 0.1% formic acid in water
* Solvent B: 0.1% formic acid in methanol
* Column temperature: 55oC
* Flow rate 0.45 ml/min

Same chromatography is used for both positive and negative mode (Table 3). | Table 3. LC gradient timetable

|  |  |
| --- | --- |
| Time | Solvent composition |
| 0 min | 98%A : 2%B |
| 20 min | 25%A : 75%B |
| 22 min | 2%A : 98%B |
| 30 min | 2%A : 98%B |
| 30.1 min | 98%A : 2%B |
| 37 min | 98%A : 2%B |

 |

### Mass spectroscopy

Agilent Technologies 6530 Accurate-Mass Q-TOF with a dual ASJ ESI ion source was used as the mass detector. Mass spectrometer settings were as follows: Ion source: gas temperature - 325 ℃, drying gas flow - 10 l/min, nebulizer pressure - 45 psig, sheath gas temperature - 400 ℃, sheath gas flow - 12 l/ml, capillary voltage - 4000 V. fragmentor voltage - 140 V, skimmer voltage - 65 V, mass range 50-1000 m/z, acquisition rate 2 spectra/s. Inline mass calibration was performed using debrisoquine sulfate (m/z 176.1182) and HP-0921 from Agilent (m/z 922.0098) in positive mode and 4-NBA (m/z 166.0146) and HP-0921 from Agilent (m/z 966.0007, formate adduct) in negative mode.

## Data analysis

Raw data processing was done using Agilent software (MassHunter Qual and ProFinder). Data analysis was performed with Agilent MassProfiler Pro package using recursive analysis workflow. Boxplots were generated using R.