Supplemental Information For:

International Ring Trial of a High Resolution Targeted Metabolomics and Lipidomics Platform for Serum and Plasma Analysis

J. Will Thompson[1, 2]*, Kendra J. Adams[1], Jerzy Adamski[3,4,5,6], Yasmin Asad[7], David Borts[8,9], John A. Bowden[10, 11], Gregory Byram[12], Viet Dang[8], Warwick B. Dunn[13], Facundo Fernandez[14], Oliver Fiehn[12], David A. Gaul[14], Andreas FR Hühmer[9], Anastasia Kalli[9], Therese Koal[15], Stormy Koeniger[16], Rupasri Mandal[17], Florian Meier[18], Fuad J. Naser[19], Donna O'Neil[13], Akos Pal[7], Gary J. Patti[19], Hai Pham-Tuan[15], Cornelia Prehn[3], Florence I. Raynaud[7], Tong Shen[12], Andrew D. Southam[13], Lisa St. John-Williams[1], Karolina Sulek[20], Catherine G. Vasilopoulou[18], Mark Viant[13], Catherine L. Winder[13], David Wishart[17], Lun Zhang[17], Jiamin Zheng[17], M. Arthur Moseley[1]

*corresponding author: will.thompson@duke.edu

1. Duke Proteomics and Metabolomics Shared Resource, Duke School of Medicine, 701 W Main Street, Durham, NC 27701; 2. Department of Pharmacology and Cancer Biology, Duke School of Medicine, Durham, NC; 3. Research Unit Molecular Endocrinology and Metabolism, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; 4. Lehrstuhl für Experimentelle Genetik, Technische Universität München, 85350 Freising-Weihenstephan, Germany; 5. German Center for Diabetes Research (DZD), 85764 München-Neuherberg, Germany; 6. Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; 7. Drug Metabolism Pharmacokinetics and Metabolomics group, Cancer Research UK Cancer Therapeutics Unit, The Institute for Cancer Research, 15 Cotswold Road, Sutton Surrey, SM2 5NG, UK; 8. College of Veterinary Medicine, Iowa State University, Ames, IA 50011; 9. Thermo Fisher Scientific, San Jose, CA; 10. National Institute of Standards and Technology, Hollings Marine Laboratory, 331 Fort Johnson Road, Charleston, SC 29412, United States; 11. Center for Environmental and Human Toxicology, Department of Physiological Sciences, College of Veterinary Medicine, 1333 Center Road, University of Florida, Gainesville, Florida 32610, United States; 12. UC Davis Genome Center – Metabolomics, Davis, CA 95618; 13. University of Birmingham and Phenome Centre Birmingham, UK; 14. School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332-0400; 15. Biocrates Life Sciences, Innsbruck, Austria; 16. AbbVie, Inc. Lake Bluff, IL; 17. Department of Computing Science, Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E8; 18. Max Planck Institute of Biochemistry, Munich, Germany; 19. Washington University, Departments of Chemistry, Genetics, and Medicine. Saint Louis, Missouri 63110 USA; 20. Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Denmark

The supplemental materials reported herein include the following documents: Ring Trial Guidance Document distributed to each laboratory for the purposes of performing the ring trial, Supplemental Table Descriptions, Supplemental Figures, and Supplemental Figure descriptions.

[Comments on the Ring Trial Guidance (added during publication): At the time the ring trial was performed, the vendor (Biocrates) was not providing explicit tune files and MS methods for each Q Exactive line instrument (i.e. classic, Plus, and HF) that had been independently validated in each lab. Because of this, each lab needed to manually program their own tune files and perform some small instrument-specific changes to the data acquisition files, following the SOP. The recommendation (not requirement) that each lab have the method checked by either Biocrates or Thermo before data acquisition was simply a safeguard against inadvertently collecting the data with a method or tune parameter programmed incorrectly. ***Currently, with the p400HR platform there is no need for vendor intervention to replicate the results, since the validated methods are provided in current versions of the kit.*** *It is of course expected that one would need to purchase the kit in order to independently replicate results.]*

RING TRIAL GUIDANCE document

Before Data Collection:

1. Since each laboratory must manually program the LC and MS data collection methods according to the p400HR kit instructions, there is more room for error in this portion of the workflow than with the p180 kit where the methods are distributed with the kit, ready to use. For this reason, it is *highly recommended* that each lab have the data collection methods checked by either Biocrates or Thermo prior to data collection. Stephen Dearth at Biocrates (stephen.dearth@biocrates.com) has kindly offered to check each group's methods via TeamViewer. Please contact Stephen directly to set this up. Even if you have already collected data, it would be great to have the methods checked. If your organization does not allow you to use TeamViewer, it is recommended that the methods be sent to Anastasia Kalli at Thermo Fisher (anastasia.kalli@thermofisher.com) in order to have the methods checked.
2. We have found that the manual calibration steps are key in order to obtain quality data with this kit, and it does not always pass the first time. The manual calibration should include:
    a. First perform mass calibration with the standard Thermo LTQ Pos Ion calibration mix.
    b. Next perform the custom mass calibration as described in the p400HR kit documentation. If this calibration does not pass, DO NOT PROCEED. Dilute the FIA Test mix 20:1 into the LTQ tune solution (instead of 10:1 as recommended in the p400HR documentation), and attempt the custom calibration again. If it still does not pass, contact Hai Pham-Tuan at Biocrates (hai.pham-tuan@biocrates.com) or Anastasia Kalli at Thermo.
3. Use a new/clean ion transfer tube for data collection. Do not use the same ion transfer tube for calibration and data collection.
4. After running the FIA test mix, load the FIA test mix data into the Skyline File FIATestMix.sky, and compare peak intensities mass accuracy to the three analysis from other instruments already loaded into the file. *Note: You must use the most recent version of Skyline –daily to perform this step.* To do this:
    a. Download the FIATestMix.sky.zip file from the following link: https://discovery.genome.duke.edu/express/resources/4786/FIATestmix.sky.zip. Save to a folder of your choosing, preferably the folder that contains the raw data you would like to analyze. *DO NOT UNZIP THE FOLDER.*

b. Download and Install the latest version of Skyline –daily build from https://skyline.ms/project/home/software/Skyline/begin.view
c. Open Skyline daily, and double-click on an icon for a blank document. Use File/Open to browse to the FIATestmix.sky.zip document, and click Open.
d. Load your FIA test mix data (one or more files) into the document by using File/Import/Results, and browsing to the files.
e. Save this document, then use the File/Share command to create a *.sky.zip archive that can be uploaded to your project in the Express Repository (https://discovery.genome.duke.edu/).
f. If your Test Mix data looks drastically different in peak intensity or shape than the files that are already loaded into the document, please contact Hai Pham-Tuan at Biocrates (hai.pham-tuan@biocrates.com).

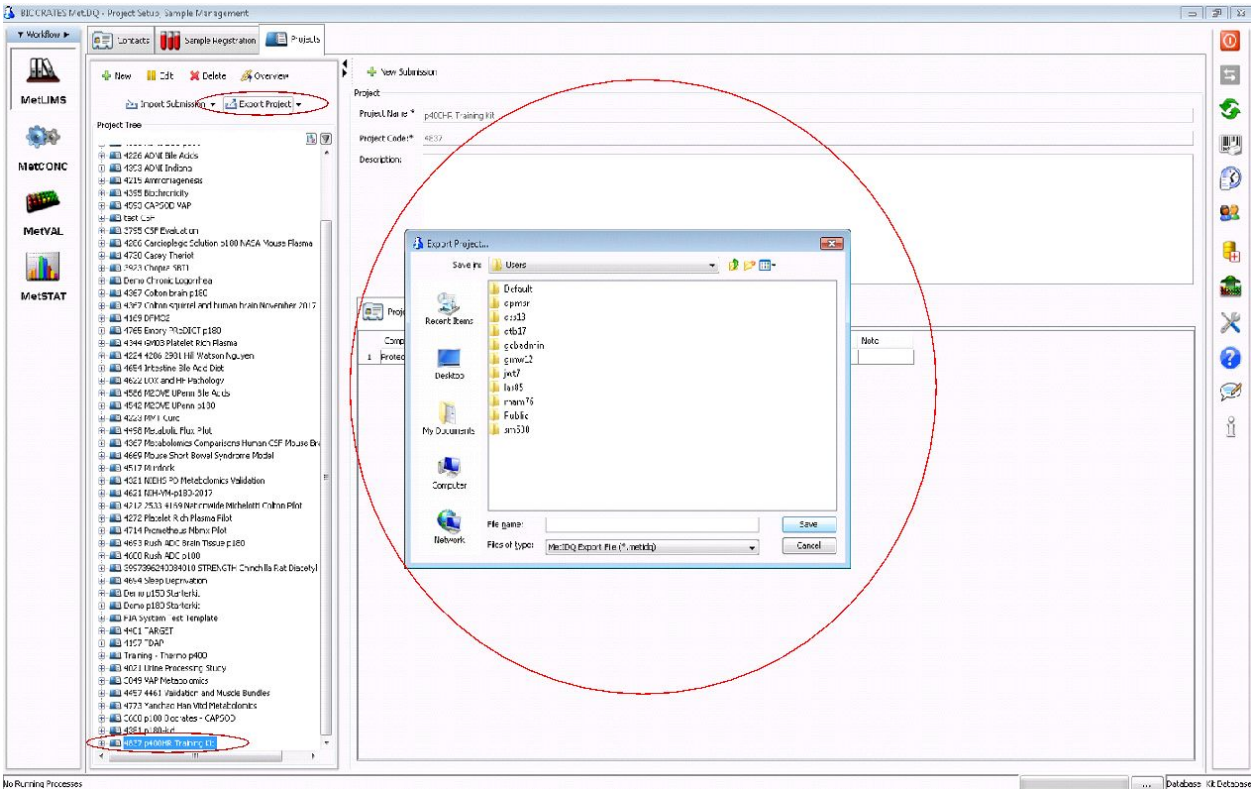
5. After running the LCMS test mix, load the FIA test mix data into the Skyline File FIATestMix.sky, and compare peak intensities mass accuracy to the three analysis from other instruments already loaded into the file. *Note: You must use the most recent version of Skyline –daily to perform this step.* To do this:
    a. Download the LCMSTestmix.sky.zip file from the following link: https://discovery.genome.duke.edu/express/resources/4786/LCMSTestmix.sky.zip. Save to a folder of your choosing, preferably the folder that contains the raw data you would like to analyze. *DO NOT UNZIP THE FOLDER.*
    b. Download and Install the latest version of Skyline –daily build from https://skyline.ms/project/home/software/Skyline/begin.view
    c. Open Skyline daily, and double-click on an icon for a blank document. Use File/Open to browse to the LCMSTestmix.sky.zip document, and click Open.
    d. Load your LCMS test mix data (one or more files) into the document by using File/Import/Results, and browsing to the files.
    e. Save this document, then use the File/Share command to create a *.sky.zip archive that can be uploaded to your project in the Express Repository (https://discovery.genome.duke.edu/).
    f. If any of the analytes in your Test Mix data look drastically different in retention time, peak intensity or peak shape than the files that are already loaded into the document, please contact Hai Pham-Tuan at Biocrates (hai.pham-tuan@biocrates.com).

After Data Collection:

1. When analyzing the LC-MS data in Quanbrowser, please follow these guidelines:
    a. Manually inspect the peak integration for each analyte and injection. The intensity column for the internal standards can be used as a visual queue for runs which may be improperly integrated, but it is best to check each injection and analyte.
    b. For each analyte, carefully inspect the calibration curves (collected in duplicate). Exclude Calibration Standards which have:
        1.>15% percent difference (inaccuracy) for Cal 2-7
        2.>20% percent difference (inaccuracy) if Cal 1

c. If you have questions about 'judgement calls' for this portion of the data analysis, please contact Will Thompson (will.thompson@duke.edu) or Lisa St.John-Williams (lisa.stjohn-williams@duke.edu) at Duke. Consistent treatment of the LCMS calibration curves between laboratories will be critical for cross-laboratory comparison.

2. Please Provide the data to Duke in two forms:
   a. Provide the Long Excel Reports for both LC Runs file that are exported from QuanBrowser (and imported into MetIDQ). Upload these to your project in Express, where XXXX is your project number (https://discovery.genome.duke.edu/).
   b. Provide the MetIDQ Project exported as XXXX.metidq as an upload to Express. The process for exporting a project from MetIDQ is shown in the screen shot below.
      1. Click on your project in the MetLIMS view in MetIDQ.
      2. Click on Export Project.
      3. Select a location in which to store your exported project so it can be uploaded to Express.



Supplemental Table Legends

Table S-1: Raw micromolar values for each sample and analyte, reported by the fourteen laboratories participating in the ring trial. Aggregate sample measures (compiled in Table 1) are calculated at the bottom of the table.

Table S-2: Inter- and Intra-laboratory variance calculations and missing data calculations for each analyte, and each sample matrix.

Table S-3: Compiled quantitative data and variance calculations for the NIST SRM-1950 plasma sample. Results from this data were further summarized in Table 2.

Table S-4: A table comparing accuracy and variance of aligned metabolites between published NIST SRM-1950 consensus values, and those obtained in the p400HR ring trial. A comparison was performed if at least n=5 laboratories reported a value in the harmonization paper of Bowden et al for the NIST SRM-1950 consensus, and the analyte had at most 20% missing values in the p400HR ring trial. Measurements from all fourteen laboratories were included in the ring trial data reported.

Supplemental Figure Legends

Figure S-1. Geographic locations of participating laboratories in the p400HR ring trial.

Figure S-2. Principal Components Analysis utilizing the retention time of 41 analytes of the UHPLC-MS system suitability test (SST) data. Two laboratories (annotated 4880 and 4904) appear to be potential outliers based on retention time, but were not excluded from subsequent analyses. "Seed" data shown in this figure is the SST data provided by Biocrates as part of the kit, not collected prior to one of the ring trial studies.

Figure S-3. Principal Components Analysis utilizing the intensity of 17 analytes of flow-injection analysis system suitability test (FIA-SST) data. One laboratory (annotated 4812) appeared to be a potential outlier due to much higher signal intensity compared to the other laboratories, but was not excluded from downstream analyses.

Figure S-4. Examples of outlier analysis using PCA for different analyte classes for all samples analyzed in the ring trial. (A) PCA for Lyso-PC lipids across all sample types shows clean separation by sample type, with to laboratory outliers. (B) The quantitative distribution plot of hexanoylcarnitine identified six analyses from laboratory 4941 with values more than 10 standard deviations higher than the others, and not all for the same sample type. (C) PCA for acylcarnitines led to the removal of the data for laboratory 4941 for the purpose of reproducibility measures in NIST SRM-1950.

Figure S-5. Principal Components Analysis examining the limits of detection (LOD) determined based on blank injections from each laboratory. Three laboratories seem to have notably different lower limits of detection than the others, and clustering analysis revealed this was because of unknown interferences in the high resolution extracted ion chromatograms. Potential laboratory, solvent, or other contamination was found to negatively impact sensitivity for some analytes.

Figure S-6. 2D Hierarchical clustering of amino acid levels using data from all laboratories and all sample types showed clear grouping by sample type and not by laboratory, supporting the thesis that with

proper system suitability and quality control, metabolomics kits such as the p400HR should enable cross-laboratory metabolomics profiling of large cohorts.

Figure S-7. Observations of one-off situational outliers which future software packages should be designed to handle. (A) A single laboratory showed consistent problems with the measurement of one analyte, tyrosine, but not with any other analytes in the UHPLC panel. (B) One laboratory had a single outlier for creatinine which was potentially due to poor peak integration.

Figure S-8. Comparison of inter-laboratory precision and accuracy between LC-MS and FIA-MS analytes. (A) Imprecision comparison for analytes with quantity >10 uM in both platforms demonstrates a T-test p-value < 0.0001. (B) Comparison of absolute value of the bias between the p400HR values and the consensus NIST SRM-1950 values, for LC platform compared to FIA platform. LC shows lower bias (T test p-value < 0.0001).

Supplemental Figures

Figure S-1. Geographic Distribution of laboratories participating in the p400HR Ring Trial.

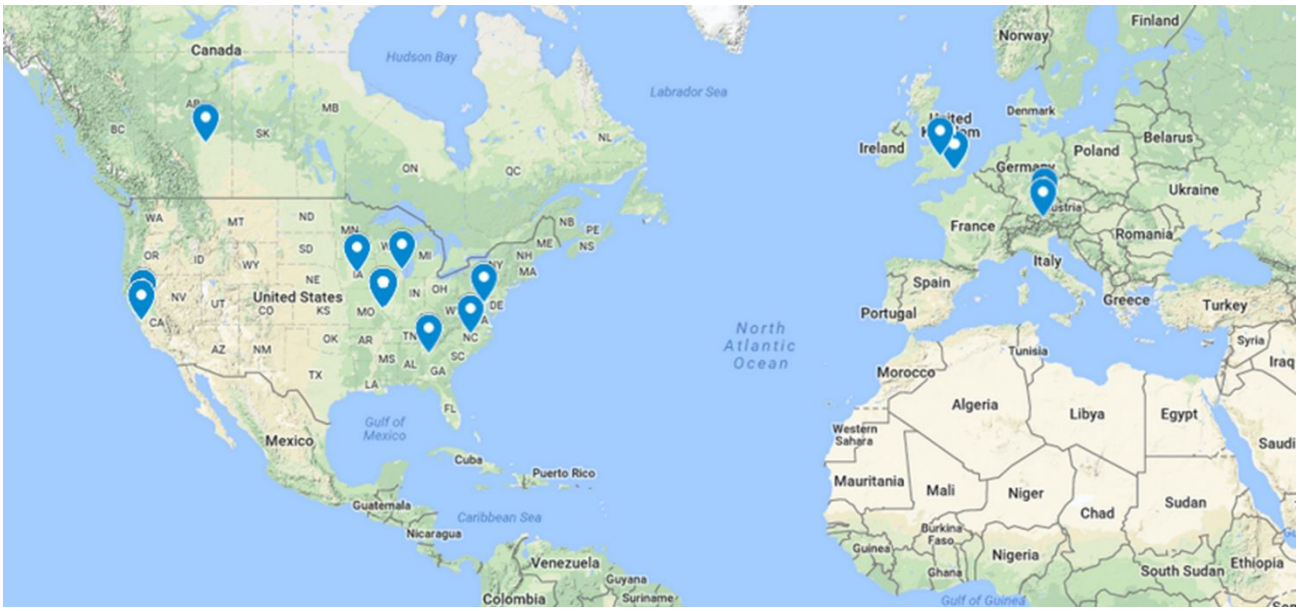


Figure S-2:

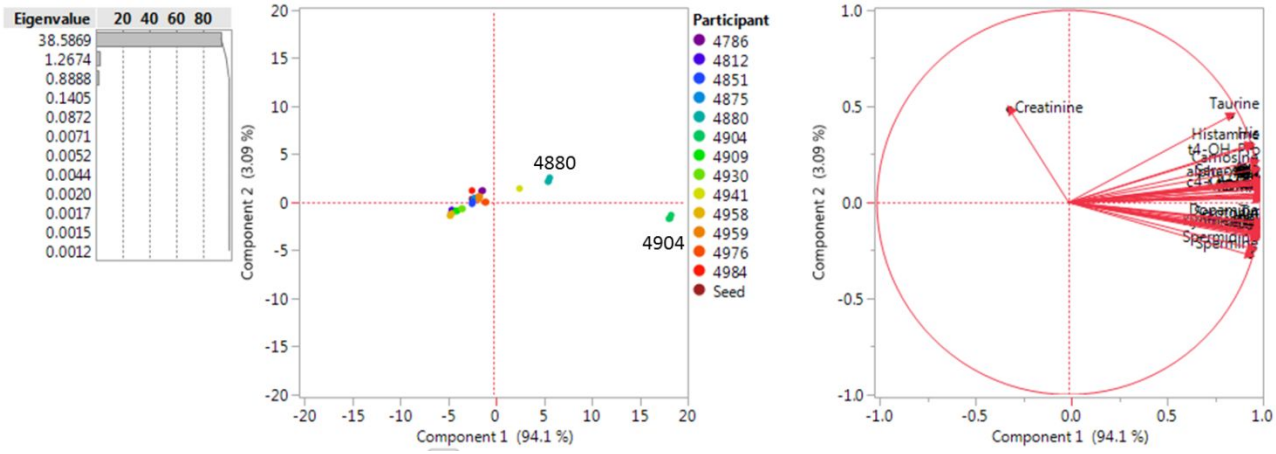


Figure S-3:

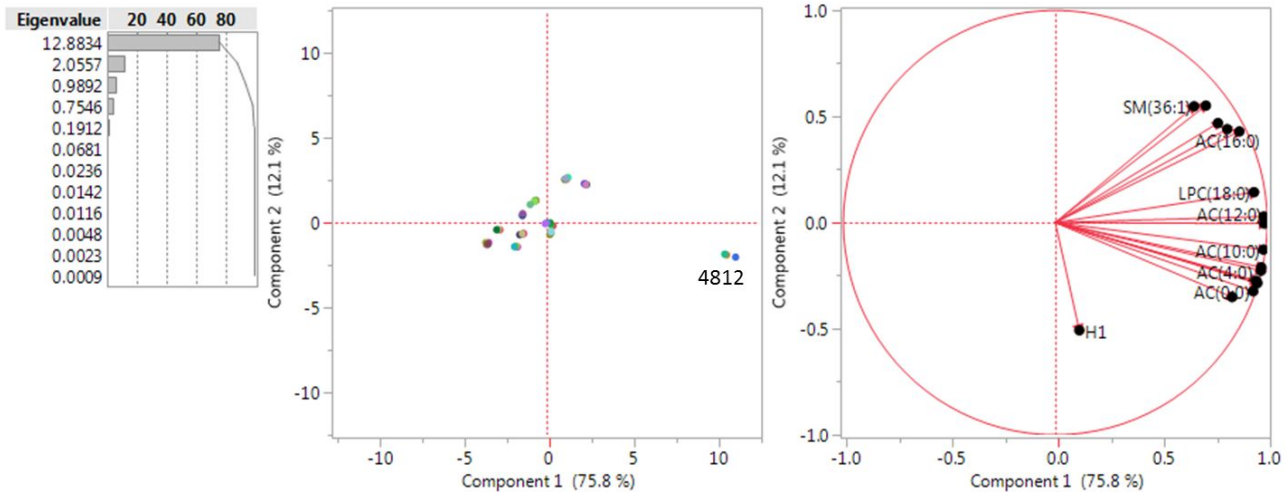


Figure S-4:

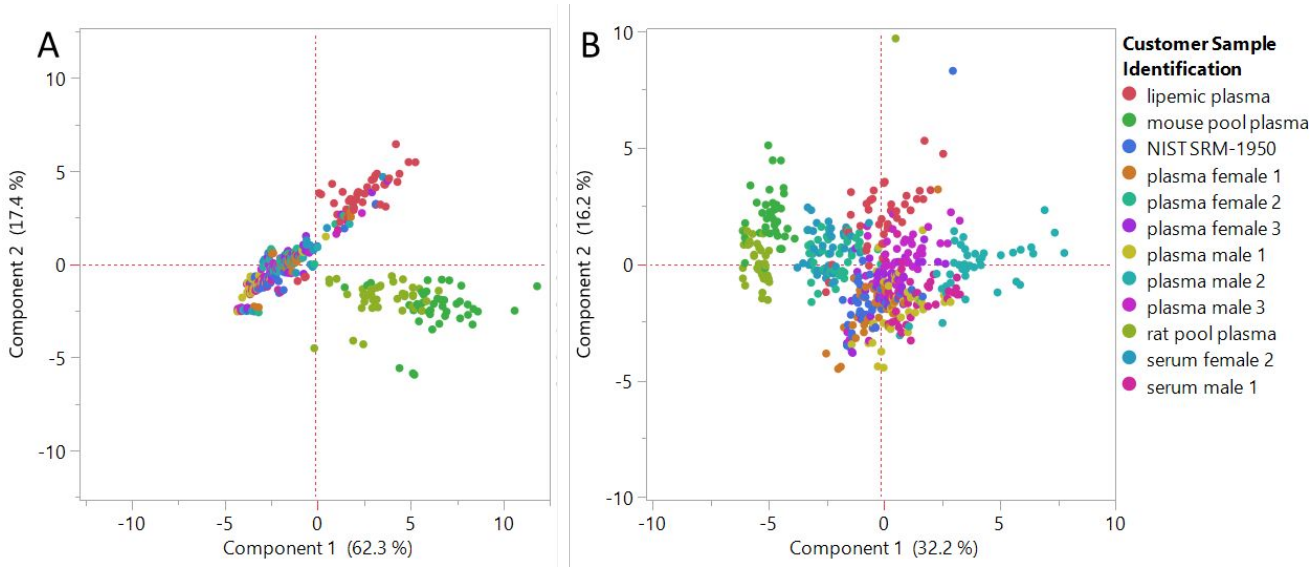


Figure S-5:

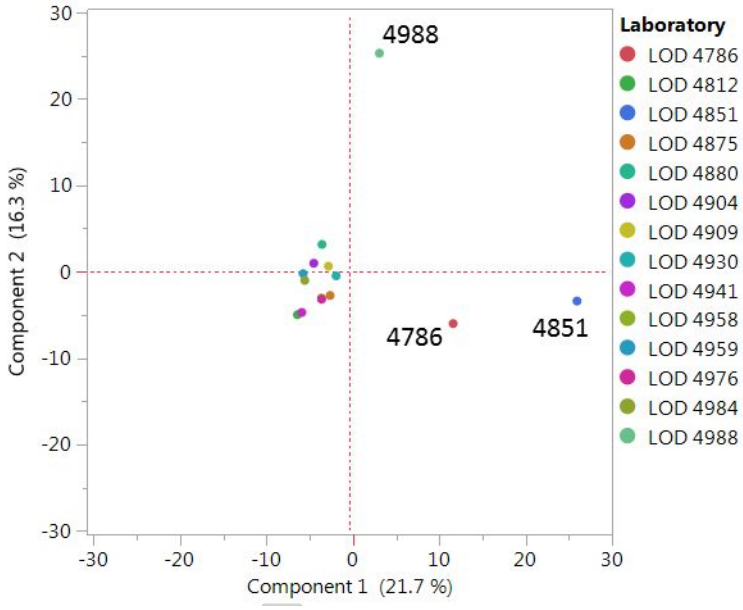


Figure S-6:

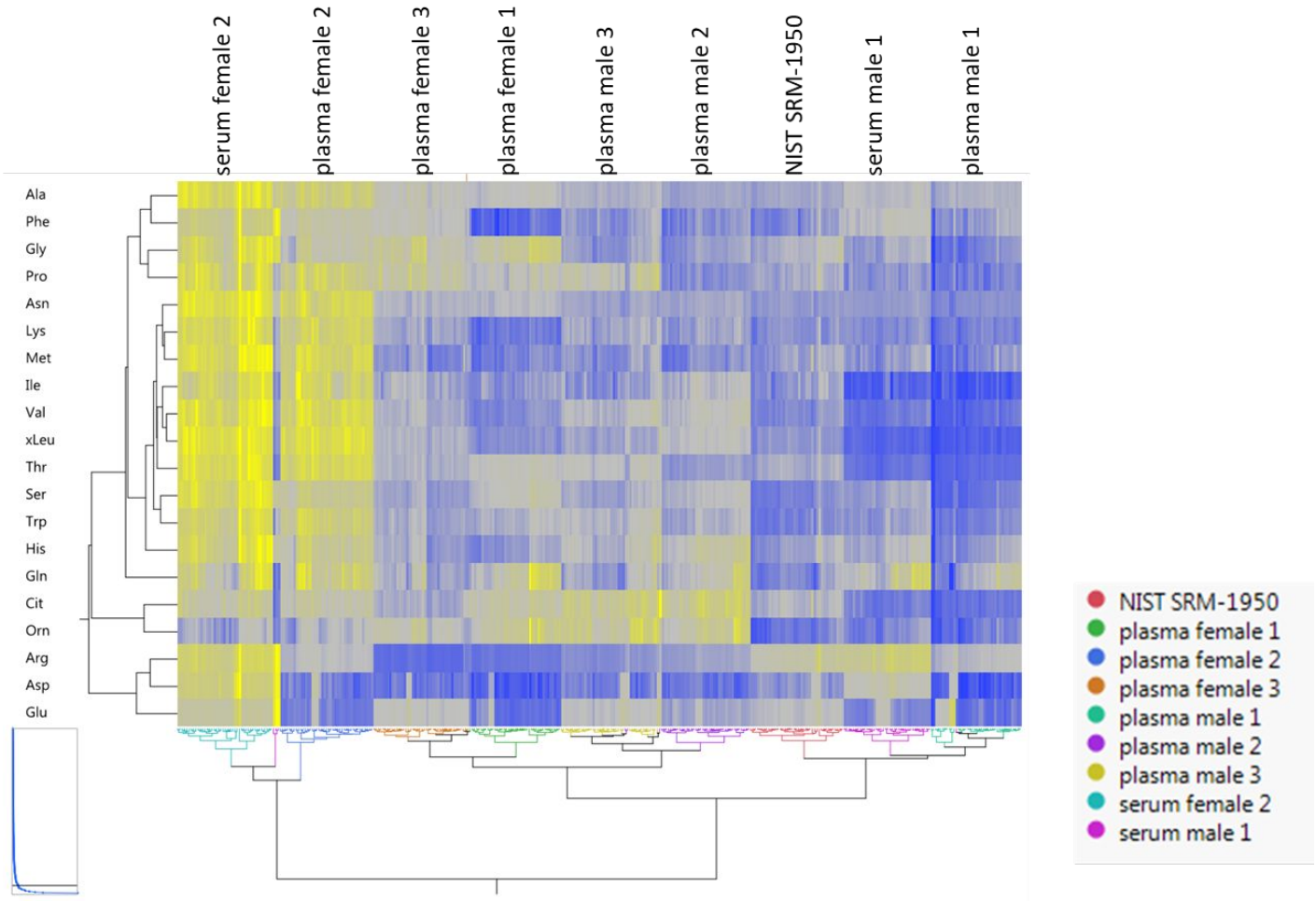

serum female 2
plasma female 2
plasma female 3
plasma female 1
plasma male 3
plasma male 2
NIST SRM-1950
serum male 1
plasma male 1
Ala
Phe
Gly
Pro
Asn
Lys
Met
Ile
Val
xLeu
Thr
Ser
Trp
His
Gln
Cit
Orn
Arg
Asp
Glu
NIST SRM-1950
plasma female 1
plasma female 2
plasma female 3
plasma male 1
plasma male 2
plasma male 3
serum female 2
serum male 1

Figure S-7:

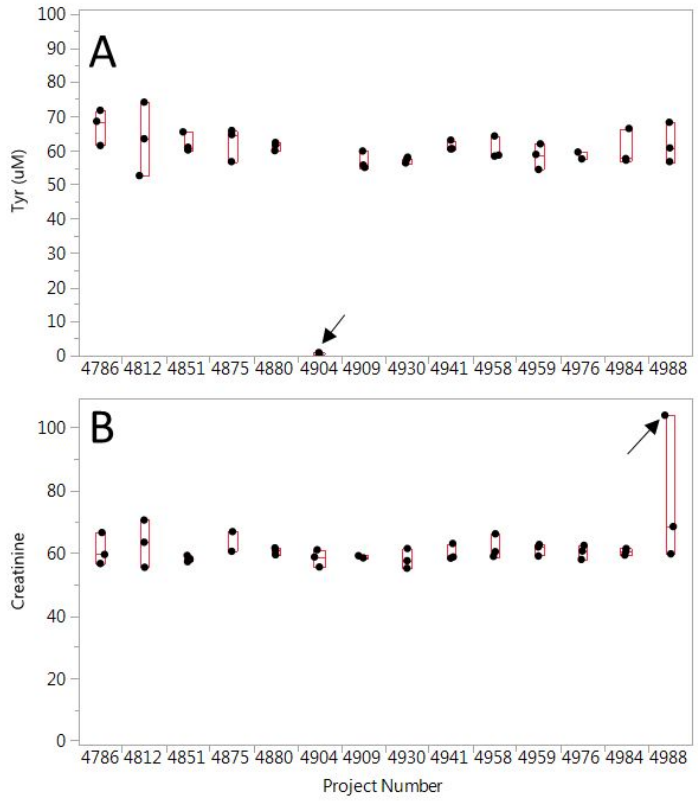


Figure S-8: