BIOMARKERS FOR SPINAL MUSCULAR ATROPHY DATA PORTAL DATA DICTIONARY

Development of the BforSMA Data Portal was a collaborative effort between the Spinal Muscular Atrophy Foundation and the Neuroscience Information Framework group at the University of California, San Diego. This work was funded by the Spinal Muscular Atrophy Foundation and the Neuroscience Information Framework (NIF) through the NIH Neuroscience Blueprint contract HHSN271200800035C via the National Institute on Drug Abuse.

DATA DICTIONARY

TABLE OF CONTENTS

Contents

Link to BforSMA Data Portal: http://neuinfo.org/bforsma/

Clinical Data

BforSMA Clinical Data

The BforSMA Clinical Data table contains subject characteristics and outcome measures results. Data from both SMA and control subjects are presented in this data table. Results can be sorted by the different outputs listed in the table or filtered by the fields in the "Filter" tab. The entire dataset can also be downloaded by going to the "Download" tab.

Outputs in data table:

- **Accession:** unique patient identifier code
- **Diagnosis:** SMA or control subject
- **SMA Type:**
  - 1: Type I defined as the inability ever to sit unaided in the judgment of the clinical investigator
  - 2: Type II defined as the ability to ever sit unaided for >30 seconds on a flat surface in the judgment of the clinical investigator
  - 3: Type III defined by the ability to ever stand unaided for 30 seconds and walk unaided for >30 feet
  - <BLANK>: control
- **Age (Years):** age in years of subject
- **Sex:** male or female
- **MHFMS Item Scores:** individual Modified Hammersmith Functional Motor Scale scores (primary outcome measure), each item is scored with a 0-2 score; 0 denotes that subject was unable to perform task, 1 denotes that subject was able to partially complete task, and 2 denotes that subject was able to completely perform task:
  - 1-Sit
  - 2-Long Sitting
  - 3-One Hand Reach
  - 4-Two Hand Reach
  - 5-Sitting to Lying
  - 6-Lifts Heads Supine
  - 7-Half Roll From Supine
  - 8-Roll Supine to Prone Over Right
  - 9-Roll Prone to Supine Over Left
  - 10-Roll Supine to Prone Over Left
  - 11-Roll Prone to Supine Over Right
  - 12-Lifts Head From Prone

  - 13-Props body on Elbows
  - 14-Props body on Extended Arms
  - 15-Body stable on Hands and Knees
  - 16-Crawls
  - 17-Lying to Sitting
  - 18-Stands Supported
  - 19-Stands Unsupported
  - 20-Stepping
- **Modified Hammersmith Functional Motor Scale Total Score:** cumulative score for the individual items of the Modified Hammersmith Functional Motor Scale; 0 minimum, 40 maximum
- **Body Mass Index Z-Score (BMI-Z):** index that is a proxy for human body fat based on subject's weight and height
  - BMI = weight (lb) / [height (in)] 2 x 703
  - BMI = weight (kg) / [height (m)]2
- **10 Meter Timed Walk Test:** a motor function test originally developed for cerebral palsy patients; a method for evaluating functional ambulation and measures time to cover 10 meters
  - NM: not measured, no assessment done either because subject was a control or subject did not qualify for the assessment (i.e. not old enough, patient too severe, etc.)
- **Current SMA Type:** classification based off of current functionality ability
  - Non-sitter: most severe SMA symptoms
  - Sitter and/or stander: less severe symptoms than non-sitters and more severe symptoms than walkers
  - Walker: least severe SMA symptoms
  - N/A: no assessment done, control subject
- **Current Level of Function:** specific classification based off of type and current functional ability
  - Type I: most severe classification
  - Type II non-rolling: 2nd most severe classification, patients cannot roll
  - Type II rolling: 3rd most severe classification, patients can roll
  - Type III non-stair climbing: 4th most severe classification, patients cannot climb stairs
  - Type III stair-climbing: least severe classification, patients can climb stars
  - N/A: no assessment done, control subject
- **Initial Age of Onset of SMA Symptoms:** number of months after birth that initial SMA symptoms were observed
  - N/A: no assessment done, control subject
- **Age SMA Diagnosed:** number of months after birth that SMA was diagnosed
  - N/A: not applicable, control subject
- **Nutritional Assessment:**

  - Modified intake/G-tube fed: subject cannot consume solid food; uses alternative method to get nutrition
  - Solid Food: subject is able to consume solid food
  - N/A: no assessment done, control subject
- **Forced Vital Capacity (Liters):** a measure of pulmonary function; performed on subjects of the mildest severity; up to 5 attempts were allowed to obtain the 3 FVC measurements that best reflect the subject's actual performance
- **Respiratory Support:**
  - BIPAP/trached: subject is on Bilevel Positive Airway Pressure (BiPAP) device or received a tracheotomy to assist with ventilation
  - Cough assist: assists with cough and airway clearance
  - NM: not measured, no assessment done either because subject was a control or subject did not qualify for the assessment (i.e. not old enough, patient too severe, etc.)

References

- Krosschell et al., A modified Hammersmith functional motor scale for use in multi-center research on spinal muscular atrophy, Neuromuscular Disorders, 2006.
- Thompson et al., Developmental Medicine & Child Neurology, 2008.
- Picture and videos of scoring MHFMS: http://smaoutcomes.org/hammersmith_manual/MHFMS_manual.php

SMN Copy Number

BforSMA SMN Copy Number

The BforSMA SMN Copy Number data table contains results and raw data/measurements used to calculate SMN1 and SMN2 copy number results. DNA samples from SMA and controls subjects were used to analyze SMN Copy Number. Results can be sorted by the different outputs listed in the table or filtered by the fields in the "Filter" tab. The entire dataset can also be downloaded by going to the "Download" tab.

Methodology

Quantification of SMN1 and SMN2 copy number was conducted at Expression Analysis Inc., Durham, North Carolina by quantitative real-time Taqman PCR (qPCR). Genomic DNA was isolated from whole blood samples using Gentra Puregene Blood Kit (Qiagen). Externally validated DNA standards for SMN1 and SMN2 were kindly provided by Dr. Wendy Chung from Columbia University Medical Center, New York, NY. The SMN1 and SMN2 reactions were carried out in 1× TaqMan Universal PCR master mix (Applied Biosystems) containing 300 nM of SMN1 primers, 250 nM of SMN probe or 450 nM of SMN2 primers, 250 nM of SMN probe, 650 nM of SMN1 non-extending oligonucleotide, respectively. The non-extending SMN1 oligonucleotide increases SMN2 assay specificity by blocking nonspecific annealing of the allele-specific primer to the opposite allele. To enable normalization of the input target DNA added to each well, the internal control CFTR gene was amplified simultaneously in a separate reaction well under identical thermal cycling conditions. The CFTR reaction was carried out wi th 1× TaqMan Universal PCR master mix containing 450 nM of CFTR primers and 250 nM of CFTR probe. PCR was performed on a 7900HT Sequence Detection System (Applied Biosystems) using a 384-well format, and amplification was achieved using the standard amplification protocol (Applied Biosystems) as follows: 50°C for 2 min, 95°C for 10 min, followed by 45 cycles of 95°C for 15 s, and 60°C for 1 min. Each reaction was run in quadruplicate with 25 ng of genomic DNA in a final volume of 15 µl. The number of SMN1 and SMN2 copies was calculated using the comparative C_T method. Results were interpreted separately by two independent investigators (Dr. Wendy Chung, Columbia University Medical Center, NY, NY; and Dr. Louise Simard, University of Manitoba, Winnipeg, Manitoba) and confirmed in a separate analysis performed by Dr. Thomas Prior at The Ohio State University using his published methodology.

References

- Gomez-Curet I, Robinson KG, Funanage VL, Crawford TO, Scavina M, et al. (2007) Robust quantification of the SMN gene copy number by real-time TaqMan PCR. Neurogenetics 8: 271-278.
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408.

Outputs in data table

- **Accession:** unique patient identifier code
- **Diagnosis:** SMA or control subject
- **Age (Years):** age in years of subject
- **Sex:** male or female
- **SMN1 Copy Number Runs (1-3):** trial run results for multiple copy number analyses
    - NM: Not Measured – no SMN1 measurements for SMA patients or repeat trial run was not necessary
- **SMN1 Copy Number Average:** average of the trial run results
    - NM: Not Measured – no SMN1 measurements for SMA patients or repeat trial run was not necessary
- **SMN1 Copy Number Final Integer:** average of the trial run results rounded to the nearest integer
    - NM: Not Measured – no SMN1 measurements for SMA patients or repeat trial run was not necessary
- **SMN2 Copy Number Runs (1-6):** trial run results for multiple copy number analyses
    - NM: Not Measured – repeat trial run was not necessary or sample analysis plate was not large enough to accommodate all samples in one trial so sample was analyzed in the next trial run
    - QNS: Quantity Not Sufficient – no results, not enough sample for analysis
- **SMN2 Copy Number Average:** average of the trial run results
    - QNS: Quantity Not Sufficient – no results, not enough sample for analysis
- **SMN2 Copy Number Final Integer:** average of the trial run results rounded to the nearest integer
    - QNS: Quantity Not Sufficient – no results, not enough sample for analysis

SMN Protein

BforSMA SMN Protein

The BforSMA SMN Protein datatable contains raw data/measurements used to calculate the SMN protein results. SMN protein was measured in peripheral blood mononuclear cells (PBMCs) using the SMN ELISA kit available at Enzo Life Sciences (http://www.enzolifesciences.com/ADI-900-209/smn-elisa-kit/). Results can be sorted by the different outputs listed in the table or filtered by the fields in the "Filter" tab. The entire dataset can also be downloaded by going to the "Download" tab. A manuscript (Kobayashi et al., PLoS ONE) with more details about the ELISA and the methodology has been submitted for publication.

Methodology

A total of 127 samples (SMA 105, control 22) were available for SMN ELISA measurements. Whole blood was collected into EDTA K_2 tubes and approximately 4mL was poured into CPT vacutainers (#362760 BD, Franklin Lakes, NJ). PBMCs were isolated by centrifugation at 1500rpm within 2 hours of collection. PBMC samples were shipped at ambient temperature to PPD (Highland Heights, KY) for further processing and frozen storage. Frozen samples were transferred to Enzo Life Sciences (Ann Arbor, MI) for SMN analysis. PBMCs were thawed in a 37°C water bath and viable cell hemocytometer counts, performed immediately prior to lysis, were used to determine the appropriate volume of lysis buffer (LB-11) needed to establish a consistent concentration of the cell suspension of 10^8 cells per milliliter. LB-11 containing 300 mM NaCl, 10% glycerol, 3 mM EDTA, 1 mM $MgCl_2$, 20 mM β-glycerophosphate, 25 mM NaF and 1% Triton X-100 was used along with protease inhibitors PIC8340 (Sigma #P8340, St. Louis, MO) and phenylmethylsulphonyl fluoride (Sigma #P7626). The cell suspension was gently vortexed and placed on ice for 30 minutes. The cell lysate was transferred to a 1.5 mL centrifuge tube and was clarified by centrifugation for 10 minutes at 14,000 RCF, 4°C. The supernatant was transferred to a clean vial and either assayed immediately or stored at -70°C until use.

Recombinant human SMN1 was generated from full-length cDNA expressed in bacterial expression vectors and purified for use as a standard in the ELISA. The capture antibody Sigma anti-SMN clone 2B1 (#S2944) was coated at 100 µL onto Costar Stripwell (#92592, Lowell, MA) at 3.5 µg/mL. After overnight incubation at room temperature, the plate was blocked for 5 hours with 1% BSA in PBS. Cell lysate samples and recombinant hSMN1 or HeLa cell lysate standards were loaded at 100 µL per well. Standards were diluted in 2-fold dilutions or from 0.0625-4 ng/mL. Samples were incubated for one hour at room temperature, washed and then incubated with a detection antibody from ProteinTech (#11708-AP-1, Chicago, IL) at 2 µg/mL for one hour at room temperature. After washing, a peroxidase conjugated goat anti-rabbit IgG from Jackson Immunolabs (#5-035-144, West Grove, PA) was applied at 50 ng/mL to the plate and incubated for 30 minutes at room temperature. After washing, plates were developed

with TMB substrate for 30 minutes at room temperature and the reaction stopped after 30 minutes with 1N HCl acid. Plates were then read on a spectrophotometer at 450 nm. Plates were sealed and gently shaken during all incubations, and dilutions of sample and standard were done in assay buffer (1% BSA, 0.1% Tween-20 in PBS).

References

- Kobayashi et al., PLoS ONE, 2011, Manuscript Submitted

Outputs in data table

- **Accession:** unique patient identifier code
- **Diagnosis:** SMA or control subject
- **SMA Type:** SMA Type I, II or III
- **Age (Years):** age in years of subject
- **Sex:** male or female
- **Assigned SMN Value (pg/mL)**: SMN protein level as per picogram per milliliters
    - <LOD: Less than Limit of Detection – no results, sample could not be measured
    - N/A: Not Applicable – no results due to QNS or <LOD
- **SMN per Cell Count (pg/10^6 PBMCs):** SMN protein level as per picogram per million PBMCs
    - N/A: Not Applicable – no results due to QNS or <LOD
- **PBMC Storage Time (Days)**
    - QNS: Quantity Not Sufficient – no results, not enough sample for analysis
- **Total PBMCs (in 10^6)**
    - QNS: Quantity Not Sufficient – no results, not enough sample for analysis
- **Estimated PBMC Viability (%)**
    - QNS: Quantity Not Sufficient – no results, not enough sample for analysis
    - <LOD: Less than Limit of Detection – no results, sample could not be measured
- **Volume of Extraction Buffer Added (mL)**
    - QNS: Quantity Not Sufficient – no results, not enough sample for analysis

SMN Transcripts

BforSMA SMN Transcripts

The BforSMA SMN Transcript data table contains raw data/measurements used to calculate the SMN transcript results. RNA samples from SMA and controls subjects were used to analyze SMN Transcripts. Results can be sorted by the different outputs listed in the table or filtered by the fields in the "Filter" tab. The entire dataset can also be downloaded by going to the "Download" tab.

Methodology

Four specific transcripts were measured: (1) SMN2-Full Length (SMN2-FL), (2) SMN1-Full Length (SMN1-FL), (3) SMN transcript lacking exon 7 (SMN-Δ7), and (4) GAPDH, a housekeeping gene commonly used for normalization in SMN transcript analysis. Three combinations of these primary measurements were also considered of potential relevance: (5) Total SMN-FL (SMN1-FL + SMN2-FL), (6) Total SMN (SMN-FL + SMN-Δ7), and (7) a ratio of SMN2 transcripts that reflects exon 7 inclusion (SMN2-FL/SMN-Δ7). As SMN1-FL is only found in control subjects, the values of SMN2-FL and Total SMN-FL are the same in SMA subjects.

SMN2-FL and SMN1-FL levels in whole blood were evaluated by absolute real-time PCR. GAPDH transcript levels were determined as positive controls both for reverse transcriptase PCR (RT-PCR) and real-time PCR in similar fashion. Analysis of SMN-Δ7 transcripts followed a procedure similar to that of SMN1-FL and SMN2-FL transcript quantification using primers SMN-Δ7-absF CTG ATG CTT TGG GAA GTA TGT TAA TT and SMN-Δ7-absR CCA GCA TTT CCA TAT AAT AGC CAG TA, and probe SMN-Δ7_absP 5'FAM - CAT GGT ACA TGA GTG GCT A -NFQ3'.

References

- Tiziano FD, Pinto AM, Fiori S, Lomastro R, Messina S, et al. (2010) SMN transcript levels in leukocytes of SMA patients determined by absolute real-time PCR. Eur J Hum Genet 18: 52-58.

Outputs in data table

- **Accession:** unique patient identifier code
- **Diagnosis:** SMA or control subject
- **SMA Type:** SMA Type I, II or III
- **Age (Years):** age in years of subject
- **Sex:** male or female
- **Read Concentration:** Amount of total transcript measured by spectrophotometer (units: ng/ul)
    - QNS: Quantity Not Sufficient – no results, not enough sample for analysis
- **SMN1-FL:** Measured SMN1 Full Length Transcript (units: mol/ng)

  - N/A: Not Applicable – no results, SMA patients do not have SMN1 transcripts
  - QNS: Quantity Not Sufficient – no results, not enough sample for analysis
- **SMN2-FL:** Measured SMN2 Full Length Transcript (units: mol/ng)
  - QNS: Quantity Not Sufficient – no results, not enough sample for analysis
- **Total SMN-FL (SMN1-FL + SMN2-FL):** Calculated Total SMN Full Length Transcript by summing Measured SMN1 Full Length Transcript and Measured SMN2 Full Length Transcript (units: mol/ng)
  - QNS: Quantity Not Sufficient – no results, not enough sample for analysis
- **SMN-Δ7:** Measured truncated delta7 SMN transcript (units: mol/ng)
  - QNS: Quantity Not Sufficient – no results, not enough sample for analysis
- **Total SMN (SMN-FL + SMN-Δ7):** Calculated Total SMN Transcripts by summing Calculated SMN Full Length Transcript and Measured truncated delta7 SMN transcript (units: mol/ng)
  - QNS: Quantity Not Sufficient – no results, not enough sample for analysis
- **SMN2-FL/SMN-Δ7:** Ratio of Measured SMN2 Full Length Transcript and Measured truncated delta7 SMN transcript (units: mol/ng)
  - QNS: Quantity Not Sufficient – no results, not enough sample for analysis
- **GAPDH:** Measured GAPDH levels, experimental control for normalization (units: mol/ng)
  - QNS: Quantity Not Sufficient – no results, not enough sample for analysis
  - <BLANK>: no results, not measured

Medications For Each Subject

BforSMA Medications for Each Subject

The BforSMA Medications for Each Subject data table contains medications for each BforSMA subject in the study. Results can be sorted by the different outputs listed in the table or filtered by the fields in the "Filter" tab. The entire dataset can also be downloaded by going to the "Download" tab.

Accession: unique patient identifier code

Specific Types of Medications: analgesic, analgesic and antihistamine, antihistamine, anti-inflammatory, chemical antacid, CNS stimulant, fatty acid metabolism, fecal softener, H2 blockers, immunizing antibody, inhaled bronchodilator, inhaled corticosteroid, lung secretions, nasal decongestant expectorant, natural bacteria, ophthalmic antibiotic, oral antibiotic, oral bronchodilator, proton pump inhibitor, putative SMN upregulator (e.g. valproic acid, phenylbutyrate, hydroxyuera), topical antibiotic, topical corticosteroid, vitamins and supplements

Frequency Options:

- Not on
- Within 14 days
- Within 30 days
- Ongoing

Medications, Non-Prescription

BforSMA Medications, Non-Prescription

The BforSMA Medications, Non-Prescription data table contains **non-prescribed** medications that BforSMA subjects were on for treatment of such as cough, acid reflux, asthma, constipation, respiratory infection, allergies, cold, respiratory congestion, etc. Results can be sorted by the different outputs listed in the table or filtered by the fields in the "Filter" tab. The entire dataset can also be downloaded by going to the "Download" tab.

Outputs in data table

- **Accession:** unique subject identifier code
- **Medication Name:** specific brand or name of the medication taken
- **Medication Ongoing:**
  - Yes: subject continued to be on the medication during the study
  - No: subject had taken the medication before was not currently on it during the study
- **Dose:** amount or quantity of medication taken
- **Unit:** standard measurement, physical quantity
- **Indication:** approved use for the drug
- **Start Date:** first date medication was taken, expressed in MM/DD/YYYY
- **Stop Date:** last date when medication was taken, expressed in MM/DD/YYYY
  - 02/02/0202 denotes that the subject is still on the medication
- **Route*:** oral (1-PO), subcutaneous (2-SC), intramuscular (3-IM), intravenous (4-IV), topical (5-TOP), inhalant (6-Inhalant), nasal (7-Nasal), nasogastric (8-NG), G-tube (9-GT), Other (99-Other)
- **Other Route*:** other self-reported route or none (-2)
- **Frequency*:** QD (1), BID (2), TID (3), QID (4), QOD (5), PRN (6), Other (99)
- **Other Frequency*:** other self-reported frequency or none (-2)

*case report form used code numbers to abbreviate the route and frequency

Medications, Other Prescription

BforSMA Medications, Other Prescription

The BforSMA Medications, Other Prescription data table contains **prescribed** medications by medical professionals that BforSMA subjects were on for treatment of cough, acid reflux, asthma, constipation, respiratory infection, allergies, cold, congestion, etc. Results can be sorted by the different outputs listed in the table or filtered by the fields in the "Filter" tab. The entire dataset can also be downloaded by going to the "Download" tab.

Outputs in data table

- **Accession:** unique subject identifier code
- **Medication Name:** specific brand or name of the medication taken
- **Medication Ongoing:**
  - Yes: subject continued to be on the medication during the study
  - No: subject had taken the medication before was not currently on it during the study
- **Dose:** amount or quantity of medication taken
- **Unit:** standard measurement, physical quantity
- **Indication:** reason for taking medication
- **Start Date:** first date medication was taken, expressed in MM/DD/YYYY
- **Stop Date:** last date when medication was taken, expressed in MM/DD/YYYY
  - 02/02/0202 denotes that the subject is still on the medication, ongoing
- **Route*:** oral (1-PO), subcutaneous (2-SC), intramuscular (3-IM), intravenous (4-IV), topical (5-TOP), inhalant (6-Inhalant), nasal (7-Nasal), nasogastric (8-NG), G-tube (9-GT), Other (99-Other)
- **Other Route*:** other self-reported route or none (-2)
- **Frequency*:** QD (1), BID (2), TID (3), QID (4), QOD (5), PRN (6), Other (99)
- **Other Frequency*:** other self-reported frequency or none (-2)

*case report form used code numbers to abbreviate the route and frequency

Metabolomics Plasma Amino Acids

BforSMA Metabolomics Plasma Amino Acids

The BforSMA Metabolomics Plasma Amino Acids data table contains raw data/measurements used from this analysis. Plasma samples from SMA and controls subjects were used the analysis. The entire dataset can also be downloaded by going to the "Download" tab.

Methodology

Metabolomics profiling was conducted on organic extracts of plasma samples using multiple analytical platforms. Amino acid analysis (AAA) was carried out as a targeted analysis of 42 species using Multiple Reaction Monitoring (MRM) on a 4000Qtrap instrument (MDS/SCIEX, Concord, ON, Canada). The amino acid analysis (AAA) platform relied on calibrated internal standards for each of the target compounds thereby facilitating absolute quantification.

To facilitate the generation of reproducible data, analytical runs were organized into batches. The batch size is defined by the number of samples that can be comfortably prepared for mass spectrometric analysis in less than a day. To be able to correct for batch-to-batch variations and to identify within-batch drifts in the analytical performance, technical replicates of a Quality Control (QC) sample are inserted at a regular interval into the run sequence. These QC samples were prepared from pooling an equal aliquot from each primary sample in the study. The QC samples are processed identically to the primary samples and results for the QC samples are used to monitor data quality. The order of acquisition of the primary samples was determined by a randomization scheme to minimize the occurrence and effect of systematic variations in/on workflow performance.

The AAA platform targeted 42 L-amino acids (including all essential amino acids). Methanol extracts of the plasma samples (10 μL) were labeled with an iTRAQ reagent (Applied Biosystems, Foster City, CA, USA) producing the *m/z* 115 reporter fragment. Known concentrations of amino acid standard labeled with the *m/z* 114 reagent were added to the sample and analyzed by LC-MRM. For each amino acid target to transitions were monitored: MH^+ -> 114 for the internal standard and MH^+ -> 115 for the unknown. The intensity ratio of the peaks, *m/z* 115 / *m/z* 114, scaled with the known concentration of the standard yielded the amino acid concentrations in μM units.

Amino Acid Analysis (AAA) data were generated for all 129 primary samples as well as 24 reference samples (153 samples in total). Samples were processed in three batches. Peak integration was performed using MultiQuant (MDS/SCIEX) software. The concentration of each detected amino acid was calculated using the in-house data processing pipeline.

Outputs in data table

- **Accession:** unique identifier associated with each patient
- **Plasma Amino Acids measured (units: μM):** Ethanolamine, Glycine, Sarcosine, β-alanine, L-alanine, β-aminoisobutyric acid, L-α-amino-n-butyric acid, L-serine, L-proline, L-valine, L-Threonine, Taurine, Hydroxy-L-proline, L-isoleucine, L-leucine, L-asparagine, L-aspartic acid, L-homocysteine, O-phosphoethanolamine, L-glutamine, L-glutamic acid, L-methionine, L-histidine, L- α-aminoadipic acid, L-phenylalanine, 1-methyl-L-histidine, 3-methyl-L-histidine, L-arginine, L-citrulline, L-tyrosine, O-phospho-L-serine, L-homocitrulline, L-tryptophan, L-ornithine, L-lysine, d-hydroxylysine, L-cystine

Metabolomics Free Fatty Acids

BforSMA Metabolomics Free Fatty Acids

The BforSMA Metabolomics Free Fatty Acids data table contains raw data/measurements for this analysis. Plasma samples from SMA and controls subjects were used the analysis. The entire dataset can also be downloaded by going to the "Download" tab.

Methodology

Metabolomics profiling was conducted on organic extracts of plasma samples using multiple analytical platforms. Free fatty acid (FFA) analysis targeted 57 free – unesterified – fatty acids by GCMS following the methylation of these compounds. The free fatty acid (FFA) platform relied on calibrated internal standards for each of the target compounds thereby facilitating absolute quantification.

To facilitate the generation of reproducible data, analytical runs were organized into batches. The batch size is defined by the number of samples that can be comfortably prepared for mass spectrometric analysis in less than a day. To be able to correct for batch-to-batch variations and to identify within-batch drifts in the analytical performance, technical replicates of a Quality Control (QC) sample are inserted at a regular interval into the run sequence. These QC samples were prepared from pooling an equal aliquot from each primary sample in the study. The QC samples are processed identically to the primary samples and results for the QC samples are used to monitor data quality. The order of acquisition of the primary samples was determined by a randomization scheme to minimize the occurrence and effect of systematic variations in/on workflow performance.

The FFA platforms uses isotope labeled fatty acid standards for absolute quantification. 41 out of the 57 targeted compounds reported by the method have their own internal standard and 16 additional fatty acids are determined against a calibration curve shared with another compound. In this platform fatty acids are converted into their form methyl esters and analyzed by GC/MS.

Free fatty acids – fatty acids not esterified to lipids – were measured by GC/MS of the corresponding fatty acyl methylesters in a format of isotope dilution measurements. 41 of the 57 target compounds were calibrated against their own isotope labeled internal standards and the remaining analytes were calibrated against one of the 48 standards.

Free fatty acid (FFA) analysis was obtained on 129 primary samples as well as 10 reference or QC samples. The project was completed in two batches. Because of the chemical stability of these compounds including the internal standards, batch correction was not necessary for this platform and

the relatively small number of QC samples was utilized only for assessing the reproducibility of the measurements.

Outputs in data table

- **Accession:** unique identifier associated with each patient
- **Plasma Free Fatty Acids measured (units: µM):** 12Me-C13:0 (iso-myristic), 12Me-C14:0 (anteiso), 13Me-C14:0 (iso), 14Me-C15:0 (iso), 15Me-C16:0 (iso), 17Me-C18:0 (iso), 18Me-C19:0 (iso), C10:0 (capric), C11:0 (undecanoic), C12:0 (lauric), C13:0 (tridecanoic), C14:0 (myristic), C15:0 (pentadecanoic), C16:0 (palmitic), C17:0 (heptadecanoic), C18:0 (stearic), C19:0 (nonadecanoic), C20:0 (eicosanoic), C22:0 (behenic), C24:0 (lignoceric), C6:0 (caproic), C7:0 (enanthic), C8:0 (caprylic), C9:0 (pelargic), c-C14:1w5 (myristoleic), c-C16:1w10 (sapienic), c-C16:1w5 (c-11-hexadecenoic), c-C16:1w7 (palmitoleic), c-C17:1w8 (c-11-heptadecenoic), c-C18:1w6 (c-12-octadecenoic), c-C18:1w7 (vaccenic), c-C18:1w9 (oleic), c-C18:2w6 (linoleic), c-C18:3w3 (linolenic), c-C18:3w6 (gamma-linolenic), c-C18:4w3 (stearidonic), c-C20:1w9 (c-11-eicosenoic), c-C20:2w6 (c-11,14-eicosadinoic), c-C20:3w3 (c-11,14,17-C20:3), c-C20:3w6 (c-8,11,14-C20:3), c-C20:4w6 (arachidonic), c-C20:5w3 (EPA), c-C22:4w6 (adrenic), c-C22:5w3 (DPA), cis-C22:6w3 (DHA, cervonic), t-C18:1w9 (elaidic), unknown (C19:2?)
    - **Unknown:** analyte could not be identified; there was low confidence of predicting correct analyte based on the acquired structural information

Metabolomics Plasma Gas Chromatography-Mass Spectrometry

BforSMA Metabolomics Free Fatty Acids

The BforSMA Metabolomics Plasma Gas Chromatography-Mass Spectrometry data table contains raw data/measurements for this analysis. Plasma samples from SMA and controls subjects were used the analysis. The entire dataset can also be downloaded by going to the "Download" tab.

Method

GC/MS of plasma samples was completed in a semi-targeted fashion: a list of analyte targets was created from a preliminary profiling experiment. The target peaks were then measured in all of the study samples. This platform used an Agilent single quadrupole mass analyzer (Model 5975, Agilent, Santa Clara, CA). The GC/MS platforms used a set of universal internal standards (5-8 non-endogenous compounds) for quantification.

To facilitate the generation of reproducible data, analytical runs were organized into batches. The batch size is defined by the number of samples that can be comfortably prepared for mass spectrometric analysis in less than a day. To be able to correct for batch-to-batch variations and to identify within-batch drifts in the analytical performance, technical replicates of a Quality Control (QC) sample are inserted at a regular interval into the run sequence. These QC samples were prepared from pooling an equal aliquot from each primary sample in the study. The QC samples are processed identically to the primary samples and results for the QC samples are used to monitor data quality. The order of acquisition of the primary samples was determined by a randomization scheme to minimize the occurrence and effect of systematic variations in/on workflow performance.

For GC/MS analysis, plasma samples (from 50 uL urine) were extracted with methanol. Dried extracts were trimethylsilylated (TMS of hydroxyl-, carboxyl-, and amino-functionalities), and oximated (oxo-functionalities), in order to make analytes volatile for GC separation. To improve analytical precision the injections were made in duplicate and the mean intensities were recorded. 128 primary samples and 54 QC samples were processed in 9 batches.

Due to the inherent complexity involved in analyte assignment to peaks identified in GC/MS, a target peak list, was carried through the analysis. Peak intensities were normalized to one of the 8 internal standards added to each sample. The most appropriate internal standard for each analyte was based on determining which internal standard exhibited the lowest variability in the QC samples for that given analyte. Peak normalization and batch correction were performed by pipelined procedures. Following statistical analysis of the data to determine marker status, identification of unknown analytes was carried out based on a priority list. Stronger markers (in terms of p-values) were assigned higher

priorities than weaker markers or non-markers. Identification of unknown analytes was attempted by a combination of matching to GC/MS library spectra (in terms of retention time and fragment masses), de novo interpretation, and purchasing standard compounds and comparing fragmentation pattern. Since plasma analytes were better represented in our GC/MS library than urine analytes, a better identification rate could be achieved for plasma analytes (~70% versus ~50% in urine).

Outputs in data table

- **Accession:** unique subject identifier code
- **Analytes that were measured (units: peak ratios):** unknown P7881_uk 11, Not evaluated, Hexadecanoic-methylester, C22:6-fatty acid, 1-Methylhistidine, D-Xylose, Arabinose, D-Ribulose or D- Xylulose, Arachidonic acid, Not evaluated 1, Acetoacetate, not evaluated 2, L-Alanine, Sarcosine, Not evaluated 3, L-Asparagine, C16:0-fatty-acid, C17:0-fatty-acid, unknown P7478_UK09, C18:0-fatty-acid, 2-Hydroxybutanoic acid, L-Proline, Unidentified, L-Ornithine, L-Valine, DL-Lactic acid, Oxalic acid, L(-)-Malic-acid, L-Histidine, unknown 31944 uk 02, 2-Hydroxypiperidine, Glutamic acid-internal-amide, L-Lysine, L-Glutamine, L-Leucine, L-Isoleucine, S-methyl-L-cysteýne, N-methyl-4-hydroxyproline, Not evaluated 4, Glycine, Beta-Alanine, Not evaluated 5, L-Lysine 2, d-Mannose, L-Methionine, Nicotinamide, Not evaluated 6, Not evaluated 7, Pyruvic acid, N-acetylaminomalonic acid, Urea, D-Glyceric acid, 3-Hydroxybutanoic acid, L-Phenylalanine, unknown 31944 uk 08, 3-Methyl-2-oxovaleric acid, 4-Methyl-2-oxovaleric-acid, Unidentified 1, Indole-3-propionic acid, L-Tryptophan, L-Serine, Not evaluated 8, Not evaluated 9, Unidentified 2, Not evaluated 10, Glycerol, Aminomalonic-acid, meso-Erythritol, 1,5-Anhydro-D-Glucitol, Unidentified 3, Not evaluated 11, Myo-inositol, L-Threonine, L-cysteýne, L-Tyrosine, Not evaluated 12, Unidentified 4, Not evaluated 13, 1-Palmitoyl-sn-glycero-3-phosphocholine, 1-Linoleoyl-sn-glycero-3-phosphocholine, C10:0-fatty-acid, L-4-Hydroxyproline, 4-Deoxyglucose, Unidentified 5, L-Aspartic-acid, Iminodiacetic acid, Not evaluated 14, Not evaluated 15, Not evaluated 16, Unidentified 6, Not evaluated 17, Monomethylphosphate, not evaluated 18, Not evaluated 19, unknown P7502_UK02, Fumaric-acid, L-Glutamic acid, Succinic acid, Not evaluated 20, Not evaluated 21, C12:0-fatty-acid, Hypoxanthine, Citric-acid, unknown 31944 uk 11, C14:0-fatty-acid, Sphingomyelin (N-base: hexadecasphing-4-enine|fatty acid C16:0), Sphingomyelin (N-base: hexadecasphing-4-enine|fatty acid C18:0), Sphingomyelin (N-base: hexadecasphing-4-enine|fatty acid C20:0), Sphingomyelin (N-base: hexadecasphing-4-enine|fatty acid C22:0), Sphingomyelin (N-base: hexadecasphing-4-enine|fatty acid C24:1), Unidentified 22, 2-Ketoglutaric acid, 4-hydroxyglutamate semialdehyde, 4-hydroxyglutamate semialdehyde 2, EDTA, Erythronic acid, Threonic acid, Sphingomyelin (N-base: heptadecasphing-4-enine|fatty acid C16:0), Sphingomyelin (N-base: heptadecasphing-4-enine|fatty acid C18:0), Sphingomyeline(N-base: heptadecasphing-4-enine|fatty acid C24:0), free Phosphate, sn-Glycerol-3-phosphate, o-

Phosphorylethanolamine, myo-Inositol 1,2-cyclic phosphate, Fructose, Sphingomyelin (N-base: sphingadiene|fatty acid C16:0), Sphingomyelin (N-base: sphingadiene|fatty acid C18:0), Sphingomyelin (N-base: sphingadiene|fatty acid C20:0), Sum|Sphingomyelin(N-base:sphingadiene|fatty acid C24:0&C24:1), C16:1-fatty-acid, Sphingomyelin (N-base: sphingosine|fatty acid C16:0), Sphingomyelin (N-base: sphingosine|fatty acid C17:0), Sphingomyelin (N-base: sphingosine|fatty acid C18:0), Sphingomyelin (N-base: sphingosine|fatty acid C22:0), Sphingomyelin (N-base: sphingosine|fatty acid C23:0), SumSphingomyelin(Nb:sphingosine|f acid C24:0&C24:1), Sphingomyelin (N-base: dihydrosphingosine|fatty acid C16:0), not evaluated 23, Inositol, D-Glucose, Creatinine, C18:2-fatty-acid, C18:1-fatty-acid, Unidentified 9, 3-Phosphoglyceric acid, 1-Palmitoyl-L-alpha-lysophosphatidic acid, Unidentified 8, Sucrose, Maltose, 1-Mono-palmitoylglycerol, 1,2-Diglyceride Molweight not derivatised = 592, 1,3-Diglyceride Molweight not derivatised = 592, 1-Mono-oleoylglycerol, 1-Mono-stearoylglycerol, 1,3-Diglyceride|Molweight not derivatised = 620, 1,2-Diglyceride|Molweight not derivatised = 620, Uric acid, Pyrophosphate, Cholesterol, Vitamin E

- **Unknown:** analyte could not be identified; low confidence of predicting correct analyte based on the acquired structural information
- **Not evaluated:** analyte could not be identified; no confidence assessment was made

Metabolomics Plasma Lipid Chromatography-Mass Spectrometry

BforSMA Metabolomics Plasma Lipid Chromatography-Mass Spectrometry

The BforSMA Metabolomics Plasma Lipid Chromatography-Mass Spectrometry data table contains raw data/measurements for this analysis. Plasma samples from SMA and controls subjects were used the analysis. The entire dataset can also be downloaded by going to the "Download" tab.

Method

Plasma lipids were analyzed by LC/MS profiling on a QStar Elite Quadrupole Time-of-Flight instrument (MDS/SCIEX, Concord, ON, Canada). The lipid platform used a set of universal internal standards (5-8 non-endogenous compounds) for quantification.

To facilitate the generation of reproducible data, analytical runs were organized into batches. The batch size is defined by the number of samples that can be comfortably prepared for mass spectrometric analysis in less than a day. To be able to correct for batch-to-batch variations and to identify within-batch drifts in the analytical performance, technical replicates of a Quality Control (QC) sample are inserted at a regular interval into the run sequence. These QC samples were prepared from pooling an equal aliquot from each primary sample in the study. The QC samples are processed identically to the primary samples and results for the QC samples are used to monitor data quality. The order of acquisition of the primary samples was determined by a randomization scheme to minimize the occurrence and effect of systematic variations in/on workflow performance.

Lipid measurements were performed on all 129 primary samples in the BforSMA study as well as 24 reference (QC) samples (153 samples in total). Analysis of the 153 samples was completed in three batches. Plasma samples were extracted with a solvent of 25%:10%:65% dichloromethane:isopropanol:methanol. All samples (primary and QC) were spiked with five internal standards that were used to track platform performance and for data normalization: 14:0 LPE, 17:0 LPC, 24:0 PC, 40:0 PC, and 51:0 TG.

LC/MS profiles of the samples were processed with a set of pipelined procedures for (LC) peak detection, peak alignment, and peak family clustering to consolidate multiple ionized forms of a lipid into a single component, normalization, and batch correction. Aligned components were quantified in terms of their processed peak intensities.

Identification of lipid components was completed by a combination accurate mass-retention time matching to known lipid species characterized previously at BGM and by LC-MS/MS analysis of the lipid

extracts. Additionally, adduct pattern and nitrogen-rule were utilized to establish the unambiguous identity of the detected lipids.

Outputs in data table

- **Accession:** unique subject identifier code
- **Analytes that were measured (units: peak ratios):** 16:0 (minor isomer) LPC, 16:0 (major isomer) LPC, 18:2 (minor isomer) LPC, 18:2 (major isomer) LPC, 18:1 (minor isomer) LPC, 18:0 (minor isomer) LPC, 18:0 (major isomer) LPC, 20:4 (major isomer) LPC, Unidentified, 16:1 CE, 16:0 CE, 18:3 CE, 18:2 CE, 18:1 CE, 20:4 CE, 20:3 CE, 16:1 SM (d18:1), 16:0 SM (d18:1), 16:0 SM (d18:0), 22:6 CE, 18:1 SM (d18:1), 18:0 SM (d18:1), 16:0/16:1 PC, 16:0/16:0 PC, 34:3 PC ae PC, 34:2 PC ae PC, 34:1 PC ae PC, 16:1/18:2 PC, 16:0/18:2 PC, 20:0 SM (d18:1), 16:0/18:1 PC, 16:1/20:4 PC ae PC, 16:0/20:4 PC ae PC, 17:0/18:2 PC, 16:0/20:5 PC, 16:0/20:4 PC, 18:1/18:2 PC, 22:1 SM (d18:1), 22:0 SM (d18:1), 18:0/18:2 PC, 18:0/18:1 PC, 18:1/20:4 PC ae PC, 18:0/20:4 PC ae PC, 23:0 SM (d18:1), 16:0/22:6 PC, 16:0/22:5 PC, 18:0/20:4 PC, 24:2 SM (d18:1), 18:0/20:3 PC, 24:1 (major isomer) SM (d18:1), 24:1 (minor isomer) SM (d18:1), 24:0 SM (d18:1), 18:2/16:0/14:0 TG, 18:1/16:0/14:0 TG, 16:0/16:0/16:0 TG, 18:0/22:6 PC, 18:0/22:5 PC, 18:2/16:1/16:0 TG, 18:2/16:0/16:0 TG, 18:1/16:0/16:0 TG, 18:0/16:0/14:0 TG, 18:3/18:2/16:0 TG, 18:2/18:2/16:0 TG, 18:2/18:1/16:0 TG, 18:1/18:0/16:0 TG, 20:4/18:2/16:0 TG, 18:2/18:2/18:1 TG, 18:2/18:1/18:1 TG, 18:1/18:1/18:1 TG, 18:1/18:1/18:0 TG, 20:4/18:1/18:1 TG
    - **Unidentified:** analyte could not be identified; low confidence of predicting correct analyte based on the acquired structural information

Metabolomics Urine Gas Chromatography-Mass Spectrometry

BforSMA Metabolomics Urine Gas Chromatography-Mass Spectrometry

The BforSMA Metabolomics Urine Gas Chromatography-Mass Spectrometry data table contains raw data/measurements for this analysis. Urine samples from SMA and controls subjects were used the analysis. The entire dataset can also be downloaded by going to the "Download" tab.

Methodology

GC/MS of urine samples was completed in a semi-targeted fashion: a list of analyte targets was created from a preliminary profiling experiment. The target peaks were then measured in all of the study samples. This platform used an Agilent single quadrupole mass analyzer (Model 5975, Agilent, Santa Clara, CA). The GC/MS platforms used a set of universal internal standards (5-8 non-endogenous compounds) for quantification.

To facilitate the generation of reproducible data, analytical runs were organized into batches. The batch size is defined by the number of samples that can be comfortably prepared for mass spectrometric analysis in less than a day. To be able to correct for batch-to-batch variations and to identify within-batch drifts in the analytical performance, technical replicates of a Quality Control (QC) sample are inserted at a regular interval into the run sequence. These QC samples were prepared from pooling an equal aliquot from each primary sample in the study. The QC samples are processed identically to the primary samples and results for the QC samples are used to monitor data quality. The order of acquisition of the primary samples was determined by a randomization scheme to minimize the occurrence and effect of systematic variations in/on workflow performance.

For GC/MS analysis, dried urine samples (from 250 µL urine) were extracted with methanol. Dried extracts were trimethylsilylated (TMS of hydroxyl-, carboxyl-, and amino-functionalities), and oximated (oxo- functionalities), in order to make analytes volatile for GC separation. To improve analytical precision the injections were made in duplicate and the mean intensities were recorded. 128 primary samples and 54 QC samples were processed in 9 batches.

Due to the inherent complexity involved in analyte assignment to peaks identified in GC/MS, a target peak list, was carried through the analysis. Peak intensities were normalized to one of the 8 internal standards added to each sample. The most appropriate internal standard for each analyte was based on determining which internal standard exhibited the lowest variability in the QC samples for that given analyte. Peak normalization and batch correction were performed by pipelined procedures. Following statistical analysis of the data to determine marker status, identification of unknown analytes was carried out based on a priority list. Stronger markers (in terms of p-values) were assigned higher

priorities than weaker markers or non-markers. Identification of unknown analytes was attempted by a combination of matching to GC/MS library spectra (in terms of retention time and fragment masses), *de novo* interpretation, and purchasing standard compounds and comparing fragmentation pattern.

Outputs in data table

- **Accession:** unique subject identifier code
- **Analytes that were measured (units: peak ratios):** Unknown, Unknown1, Not evaluated, 3-Methylhistidine, 1-Methylhistidine, Not evaluated1, Not evaluated2, Not evaluated3, 2,4-Dihydroxybutanoic acid, D-Xylose, Arabinose, Arabitol, D-Ribose, deoxyfructose or isomer, Not evaluated4, Psicose, Psicose1, 2-Mono-oleoylglycerol, hexanedioic acid, L-Alanine, L-Asparagine, fucose, 2-Hydroxyglutaric acid, N-methylaminoethanol, 2-hydroxyvaleric acid or 3 hydroxyisovaleric acid, Glycolic Acid, L-Ornithine, DL-Lactic acid, Oxalic acid, L-Histidine, Glutamic acid-internal-amide, pantoic acid, Ascorbic acid, N-Acetylaspartic acid, Not evaluated5, 3-Hydroxyphenylacetic acid, Benzylalcohol, unidentified, Not evaluated6, Not evaluated7, 4,6-dihydroxy-2-methylpyrimidine, Unidentified70, Isovalerylglycine, Beta-Alanine, Not evaluated8, d-Mannose, Not evaluated9, 3-hydroxypropionic acid, Benzoic acid, 4-Hydroxyphenylacetic acid, Unidentified1, Not evaluated10, Pyruvic acid, Not evaluated11, L-Phenylalanine, Not evaluated12, Unidentified2, L-Tryptophan, Not evaluated13, Unidentified3, Galactosylglycerol, Lactose, Not evaluated14, Not evaluated15, Not evaluated16, Mannitol, Not evaluated17, Hippuric acid, Homo-vanillic acid, Not evaluated18, N-acetylglucosamine, N-acetylgalactosamine, Not evaluated19, meso-Erythritol, 3,4,5-Trihydroxypentanoic acid, Not evaluated20, Myo-inositol, Pseudo-uridine, Not evaluated21, Not evaluated22, L-Threonine, Aminomalonic acid, N-(hydroxy-methoxybenzoyl)glycine, Not evaluated23, free Sulphate, Not , valuated24, Not evaluated25, cis-Aconitic acid, L-4-Hydroxyproline, 3-Hydroxybutanoic acid, L(-)-Malic-acid, Not Evaluated70, Unidentified4, Not evaluated26, Not evaluated27, Isocitric acid, Not evaluated28, L-Glutamic acid, 2-Methylmalic acid, 3-hydroxyglutaric acid, Unidentified5, Unidentified6, Not evaluated29, Allantoin, Not evaluated30, DL-alpha-Aminoadipic acid, Hypoxanthine, Not evaluated31, Not evaluated32, Unidentified7, 5-hydroxymethyl-2-furancarboxylic acid, Not evaluated33, Citric-acid, Not evaluated34, not evaluated71, L-Tyrosine, Unidentified8, 2-Ketoglutaric acid, 5-Hydroxyindole-3-acetic acid, Unidentified9, Erythronic acid, Threonic acid, Tartaric acid, Ribonic acid, Not evaluated35, Not evaluated36, Not evaluated37, N-(3-hydroxybenzoyl)glycine, Not evaluated38, Not evaluated39, Vanillic acid, Unidentified10, Not evaluated40, Not evaluated41, Not evaluated42, sn-Glycerol-3-phosphate, o-Phosphorylethanolamine, Not evaluated43, Not evaluated44, Fructose, Unidentified11, Inositol, D-Glucose, N-(2-hydroxybenzoyl)glycine, Not evaluated45, Not evaluated46, Creatinine, Unidentified12, Unidentified13, Ascorbic acid1, Glucuronic acid, Gluconic acid, Not evaluated47, Quinic acid, Xanthine, 6,8-dihydroxypurine, Sucrose, Unidentified14, Not evaluated48, Not

evaluated49, 1-Methyl uric acid, 1-Methyl uric acid1, D-Glucose-6-phosphate, Not evaluated50, Not evaluated51, Pantothenic acid, Uric acid, Pyrophosphate

- **Unknown:** analyte could not be identified; low confidence of predicting correct analyte based on the acquired structural information
- **Not evaluated:** analyte could not be identified; no confidence assessment was made

Nutritional Supplements and Vitamins

BforSMA Nutritional Supplements and Vitamins

The BforSMA Nutritional Supplements and Vitamins data table contains supplements and vitamins that BforSMA subjects were on before and during the study. Results can be sorted by the different outputs listed in the table or filtered by the fields in the "Filter" tab. The entire dataset can also be downloaded by going to the "Download" tab.

Outputs in data table

- **Accession:** unique subject identifier code
- **Supplement/Vitamin Name:** specific brand or name of the supplement/vitamin taken
- **Supplement Ongoing:**
  - Yes: subject continued to be on the medication during the study
  - No: subject had taken the medication before was not currently on it during the study
- **Dose:** amount or quantity of supplement/vitamin
- **Unit:** standard measurement, physical quantity
- **Indication:** reasons for taking supplement/vitamin
- **Start Date:** first date supplement/vitamin was taken, expressed in MM/DD/YYYY
- **Stop Date:** last date when supplement/vitamin was taken, expressed in MM/DD/YYYY
  - 02/02/0202 denotes that the subject is still on the supplement/vitamin, ongoing
- **Route*:** oral (1-PO), subcutaneous (2-SC), intramuscular (3-IM), intravenous (4-IV), topical (5-TOP), inhalant (6-Inhalant), nasal (7-Nasal), nasogastric (8-NG), G-tube (9-GT), Other (99-Other)
- **Other Route*:** other self-reported route or none (-2)
- **Frequency*:** QD (1), BID (2), TID (3), QID (4), QOD (5), PRN (6), Other (99)
- **Other Frequency*:** other self-reported frequency or none (-2)

Plasma Proteomics

BforSMA Plasma Proteomics

The BforSMA Plasma Proteomics data table contains raw data/measurements from the plasma protein analysis. Plasma samples from SMA and controls subjects were used the analysis. Specific proteins can be searched by using the "Protein name" autocomplete search box in the "Filter" tab. The entire dataset can also be downloaded by going to the "Download" tab.

Methodology

The quantitative discovery proteomics workflow was based on multi-dimensional liquid chromatography – MS/MS analyses of peptides combined with 8-plex iTRAQ labeling. To achieve sufficient dynamic range of plasma analysis and long-term reproducibility, a two-stage protein depletion method was optimized. In the first depletion stage, 14 abundant plasma proteins are depleted by an IgY14 antibody column (Sigma-Aldrich, St. Louis, MO, USA) (serum albumin, IgG, fibrinogen, transferrin, IgA, IgM, haptoglobin, α-2-macroglobulin, α-1-acid glycoprotein, α-1-antitrypsin, Apo A-I, Apo A-II, complement C3, and Apo B-100) The flow-through was further depleted by a Supermix column (Sigma-Aldrich), which retains moderately abundant proteins with a broad specificity. Proteins in the Supermix flow-through are recovered on a reversed-phase column.

Protein samples were reduced with TCEP, alkylated with iodoacetamide, and digested with trypsin. Following digestion each sample is labeled with a discrete iTRAQ reagent. Six of the eight channels were utilized for primary samples and two were used for reference samples created by combining aliquots from each primary sample in the study. Labeled samples were combined into an iTRAQ mix and fractionated by strong cation exchange into six fractions. Each of the fractions were further fractionated by HPLC and spotted on MALDI plates for MSMS analysis using an AB/SCIEX 4800 TOF/TOF mass spectrometer (MDS SCIEX, Concord, ON, Canada). Acquisition of LC-MSMS data was optimized by in-house developed algorithms to select and measure consistent sets of peptides from experiment to experiment.

Relative quantification of peptides was carried out by determining relative intensities of reporter ions between the sample specific channels (*m/z* 114, 115, 116, 118, 119, 121) and reference sample channels (*m/z* 113, 117). The average ratios relative to the two reference channels were used in most experiments.

Identification of peptides from the MS/MS spectra was achieved using the Mascot database searching tool and a BGM-developed validation protocol to distinguish true and false positive peptide matches. Once data collection was completed for every study sample, peptides were assigned to a minimum non-

redundant protein set. Relative quantification of proteins was achieved by assigning the median ratio from peptides mapped to the given protein. Normalization of protein expression data was carried out using a procedure described by Vandersompele et al.

References

- Juhasz P, Lynch M, Sethuraman M, Campbell J, Hines W, et al. (2011) Semi-targeted plasma proteomics discovery workflow utilizing two-stage protein depletion and off-line LC-MALDI MS/MS. J Proteome Res 10: 34-45.
- Qian WJ, Kaleta DT, Petritis BO, Jiang H, Liu T, et al. (2008) Enhanced detection of low abundance human plasma proteins using a tandem IgY12-SuperMix immunoaffinity separation strategy. Mol Cell Proteomics 7: 1963-1973.
- Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, et al. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3: 1154-1169.
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3: RESEAsRCH0034.

Outputs in data table:

- **Accession:** unique subject identifier code
- **IPIID:** international protein index code (http://www.ebi.ac.uk/IPI/IPIhelp.html)
- **Protein Name:** full name of the protein analyte
- **NCBIID:** National Center for Biotechnology Information protein ID (http://www.ncbi.nlm.nih.gov/protein)
- **Gene Name:** unique gene symbol (http://www.genenames.org/)
- **Mass/charge ratio:** *m/z* results for each protein from the proteomics analysis
- **Analytes that were measured (units: peak ratios):**

CILP2	QSOX1	COL6A1	AMBP	APP	PGD	GALNT2	FAM20C	HABP2	FCGR2A
COMP	PROC	TLN1	LBP	LOC442497	AGT	VASN	LRP1	APOA1	IGFBP4
TNXB	PGK1	THBS1	C2	APOH	LILRB5	PSMA1	IGF1	TIMP1	Unknown
CLEC3B	CST6	SELENBP1	SH3BGRL	SRGN	EXTL2	PTPRU	AMH	SIGLEC5	LCN2
ADAMTSL4	S100A4	SPINK1	FGB	MDH1	IGFBP1	TCN1	C4A	C8B	EXT1
TNXB	SERPINA10	IGFBP7	TPM4	MCAM	HLA-C	CLSTN1	C8A	FTL	TMEM132C
DPP4	ENG	Unknown	MASP1	Unknown	CRELD1	CPN1	COL11A2	BASP1	B3GAT3
THBS4	PARK7	F9	PSME2	IL6ST	GGH	ITIH1	AHSG	FCGBP	CST3
CDH13	VCAM1	AOC3	RHOXF1	PKM2	KIT	SLC38A10	FLT4	SPON1	H6PD

OMD	MET	CDH2	ITGB1	ECM1	B3GNT1	Unknown	ACAN	IGHG4	MAN1A1
CRTAC1	PEBP4	SELL	SERPING1	PTPRF	SERPINA7	BTD	ICAM2	LTBP4	C1S
F13B	ORM2	HSPG2	COL1A1	GP1BA	HPX	IGF2R	CD163	PRG2	PTGDS
CD93	CNDP1	PPIA	TTR	FUCA1	Unknown	CSK	KRT10	IGHA1	PPIB
LUM	FAP	IGKV4-1	CR2	GC	ALK	IGH@	SDC1	CDH1	CTSZ
PEPD	CFL1	ENO1	P4HB	HSPA8	FETUB	ALDOB	SEPP1	LYVE1	Unknown
APCS	VNN1	FKBP1A	MMP9	PTPRM	EPHA1	PON3	KLKB1	NME3	KRT1
COL2A1	CA1	PFN1	SAA1	HPR	UMOD	TKT	F10	ADAMTS13	FAM3C
NOV	PTPRG	EPHA4	WARS	DPEP2	C1QB	ICOSLG	CHL1	ENPP2	Unknown
Unknown	MB	YWHAZ	COL6A3	ANPEP	NAGLU	FUCA2	Unknown	ALB	IGL@
COL6A3	NID1	ALAD	SHBG	HSPA5	IGFALS	GPLD1	SPP2	C6	RNASE2
CRP	RCTPI1	Unknown	C3	GPNMB	PCOLCE	SIRPA	CD109	RTN4RL2	C1QC
VTN	PF4	Unknown	TIMP2	GSTO1	SERPINF1	ADIPOQ	Unknown	COL5A1	IGHG2
GSN	PGCP	COL1A2	CTBS	NPC2	MST1	ANPEP	APOC1	KRT9	SOD2
VTN	S100A8	COL18A1	IGJ	APOA4	MASP1	Unknown	Unknown	APOB	FCGR3B
HBB	GAPDH	NME1	FAH	GOLPH3	F11	TF	KNG1	LSAMP	PTPRJ
ALPL	CNTN4	Unknown	LILRA3	GPX3	TFRC	GRN	NRCAM	Unknown	F5
Unknown	SHBG	SERPINA6	IGFBP5	PVR	FSTL1	EFEMP1	CPN2	GOLM1	LAMP1
FCGR3A	TAGLN2	FGA	CBLN4	ARHGDIB	MMP2	HYOU1	BST1	LTBP1	HSP90B1
NEO1	AFM	ELF1	APOF	APOL1	C2	MBL2	PON1	PLXNA1	PSMB2
Unknown	NP	HBB	EIF5A	C8G	VWF	PSMC3IP	C4A	Unknown	MEGF8
LRG1	HP	ITIH2	FGA	Unknown	CRHBP	ATF6	CD44	Unknown	CD14
CTSD	DAG1	GM2A	PLXNB2	A2M	CAST	IGHD	ANG	CLPX	AMY2B
C2	PROCR	PI16	A2M	ABI3BP	ITIH4	SERPINF1	PTPRS	SERPINC1	ALCAM
RPS27A	ACTN1	GNPTG	C5	PROZ	CNDP1	PTK2B	SERPINA4	CORO1A	DKK3
CFI	CSF1R	MGP	CNTN1	DBH	TGFBI	CPB2	OSCAR	C1R	F7
IGFBP6	PRDX2	ITGA7	PLTP	VCL	IGF2	IGFBP2	FGFR1	LOC100289383	LAMB1
INHBC	GSS	S100A6	FGFR4	IL1RAP	LAMA2	PRAP1	SCARA3	ITIH3	OPCML
F13A1	S100A9	FGG	ROBO4	SILV	FBLN1	VWF	APOA2	FCN2	Unknown
NCAM1	TXN	CFD	NOTCH2	CETP	CECR1	ACTB	MRC1	PLXDC2	LYZ
CP	SOD1	B3GNT8	NRP1	F5	RNASE4	CLU	MASP2	PODXL	ENPEP
AGA	PRDX6	PZP	HEG1	FGL2	CFP	SERPINE1	LAMP2	C7	C1QTNF5
C9	SERPIND1	BPGM	HGFAC	ROBO1	IGHG3	LMAN2	PGLYRP2	MOS	PRCP
HBA1	MMRN2	PCSK9	CALR	CD248	CAST	PLD4	MINPP1	SERPINA11	A1BG
LRG1	CHAD	SH3BGRL3	Unknown	ACSM3	SERPINA7	B2M	EXT2	ICAM3	APOC2
SPP1	LCP1	CDH5	THBS1	APOA4	ICAM1	RBP4	A2M	AXL	BCHE

CKM	CACNA2D1	PPBP	SPARCL1	SERPINA1	LCAT	APOM	SERPINA5	APOC3	Unknown
ISLR	Unknown	DDT	PROS1	TIE1	RARRES2	FMOD	PTPRD	SPARC	FBLN1
PRG4	CAT	TMSL3	SAA4	POSTN	Unknown	FGL1	HAPLN1	Unknown	SEMA3F
MRC2	PDIA3	CHGA	PRKCSH	CSPG4	EFEMP2	C1RL	COL3A1	LGALS3BP	TNC
CA2	ORM1	NOTCH3	AMBP	C1RL	PNLIP	APLP1	DSG2	LDHB	CFH
ANGPTL3	CNTN3	LOC100133794	TRIM41	Unknown	MERTK	EPAS1	RNF169	TCN2	TCL6
NEGR1	C22orf31	ARPC3	GALNS	RNASE1	DNHD1	COL15A1	PDCD1LG2	MSN	GPR123
F2	POMGNT1	OIT3	PLXND1	IL1R2	C6orf174	COG7	SPTBN1	IDH3A	CCL16
MMRN1	GPD1	IRX3	NUTF2	PLG	PPIL2	PEBP1	HPCA	CFHR1	TGOLN2
CA11	STOML2	PVRL1	HEXB	OLFM1	DBI	KCNG4	CEP135	TRA@	SEMA7A
FLNB	PCNT	PSMB1	APC	FBN1	DCUN1D1	SPTAN1	PUF60	SOD3	COLEC11
SERPINF1	GUCA2A	PGAM2	PDCD6	Unknown	SEPSECS	CPD	KIF3B	IGL@	ACE
CEACAM1	KLHDC7B	RREB1	CANT1	SEC31B	LOC100289290	GPT	DCTN1	PSME1	GALNT7
APOE	BXDC1	TK1	CABC1	CADM1	VCP	NFKB2	TUBA4A	TRAF5	MKI67
CTSB	UBE2V1	PPEF1	PDIA6	Unknown	MEGF10	REG1A	COTL1	ADAMDEC1	RPN2
FCN3	TYMP	ZNF496	FABP4	F2	PAPLN	FBXO38	COBRA1	OGN	RGS12
CCDC80	NCAM2	NAALADL2	ITGAM	IGDCC4	TGFBR3	SRPRB	SMPDL3A	GOT1	FCRL5
CNTFR	GPR116	PLEKHA5	MED8	B3GNT2	ANXA3	FBLN5	B4GALT1	PAM	BLMH
ESAM	PRDX5	PIGR	APEH	IGFBP3	FCGRT	NAP1L4	RPL21	OAF	NUCB1
ATRN	C9orf141	PCGF2	LOC100289629	F12	ZNF770	NAMPT	HIC2	Unknown	CHI3L1
AZGP1	TIMD4	CDH6	FTSJD2	FN1	LAMB2	DNAH5	ANAPC1	F10	RNH1
APOD	MAN2A2	MCM10	DDX53	LTA4H	PLXDC1	CUTA	ME3	ALDOC	TALDO1
SEMA4B	SPINK5	SLC14A2	GHR	Unknown	FAM135B	KCTD12	SQLE	FERMT2	SLC25A11
Unknown	MARCO	SMCR8	NAPRT1	HRG	CENPF	LSG1	NAPA	SLC33A1	PLEKHG4B
APOC4	CD5L	NPHS1	LOC344065	ALDOA	EYA3	CALCR	ATP8B3	NPHS2	LOC100130248
									NDEL1