# Joint and Unique Multiblock Analysis of biological data – multiomics malaria study

# **Supporting Information**

Izabella Surowiec<sup>1,2</sup>, Tomas Skotare<sup>1</sup>, Rickard Sjögren<sup>1</sup>, Sandra Gouveia-Figueira<sup>1</sup>, Judy Orikiiriza<sup>3,4,5</sup>, Sven Bergström<sup>6</sup>, Johan Normark<sup>6</sup>, Johan Trygg<sup>1,2\*</sup>

- 1. Computational Life Science Cluster (CLiC), Department of Chemistry, Umeå University, Linnaeus väg 10, 901 87 Umeå, Sweden
  - 2. Sartorius Stedim Data Analytics, Tvistevägen 48, 907 36 Umeå, Sweden
- Infectious Diseases Institute, College of Health Sciences, Makerere University, P.O. Box 22418, Kampala, Uganda
- Department of Immunology, Institute of Molecular Medicine, Trinity College Dublin, St. James's Hospital, Dublin 8, Ireland
  - 5. Rwanda Military Hospital, P.O. Box: 3377, Kigali, Rwanda
  - 6. Department of Molecular Biology, Umeå University, 901 87 Umeå, Sweden

# Methods

## Sample Extraction for GCMS and LCMS Metabolomics Analysis

Frozen plasma samples were thawed at 37 °C for approximately 15 min. Low molecular weight compounds (< 900 Da) were extracted by adding 900  $\mu$ L extraction solution (a mixture of methanol and water (8:1, v/v) with internal standards) to 100  $\mu$ l plasma and vortexing for around 10 s. Next, the samples were vigorously shaken at 30 Hz for 1 min in a MM301 Vibration Mill (Retsch GmbH & Company KG), and then placed on ice for 120 min followed by centrifugation at 14000 rpm, 4°C. Aliquots of 200  $\mu$ l of the supernatant were transferred to separate vials for subsequent LC-MS and GC-MS analysis. The transferred supernatants were evaporated to dryness under vacuum using a miVac QUATTRO concentrator (Genevac LTD, Ipswich, UK) for approximately 2 h at room temperature. Vials with dried extracts were stored at -80°C until analysis.

# Sample Extraction for Lipidomics LCMS Analysis

Plasma was thawed on ice and 110  $\mu$ L of extraction mixture (chloroform:methanol (2:1, V/V) were added to 20  $\mu$ L of plasma and extraction carried out using a MM301 vibration Mill (Retsch GmbH & Co. KG, Haan, Germany) at a frequency of 30 Hz for 2 min. Then samples were stored at ambient temperature for 60 minbefore being centrifuged at 14 000 rpm for 3 min at 4°C. A 50  $\mu$ L aliquot of the resulting lower phase was transferred to a LC vial, 70  $\mu$ L of a chloroform:methanol (2:1, V/V) mixture were added and the samples shaken briefly before being analyzed by LCMS.

# GCMS Derivatization and Analysis

To each dry extract, 30  $\mu$ l of methoxyamine solution in pyridine (15 $\mu$ g/ $\mu$ L) was added and then shaken for 15 min on a shaking table. Derivatization was carried out at 70°C for 1 h followed by room temperature for 16 h. Subsequently, the samples were trimethylsilylated with 30  $\mu$ L MSTFA at room temperature for 1 h. Finally, 30  $\mu$ L of heptane (with 15 ng methylstearate/ $\mu$ L) was added and the vials were vortexed before 1  $\mu$ L was injected splitless by a CTC Combi Pal autosampler (CTC Analytics AG, Switzerland) into an Agilent 6890 GC equipped with a fused silica capillary column (10 m x 0.18 mm I.D) with a chemically bonded 0.18  $\mu$ m DB 5-MS stationary phase (J&W Scientific). The injector temperature was 270°C and the purge flowrate was 20 ml/min (the purge was turned on after 60 s). The column temperature was 70°C for 2 min, then increased by 40°C/min to 320°C, and held there for 2 min using a gas flow rate of 1 mL/min. The GC was coupled to the ion source of a Pegasus III TOF-MS (Leco Corp., St Joseph, MI, USA). An alkane series (C10-C40) was run together with all samples.

#### LCMS Metabolomics Analysis

On the day of analysis, samples were reconstituted with 20  $\mu$ L internal standard solution (methanol/water, 1:1) at a concentration of 50 ng/ $\mu$ L (phenylalanine (U- 13C<sub>9</sub>), caffeine (trimethyl-13C<sub>3</sub>), cholic acid (2,2,4,4-d<sub>4</sub>), arachidonic acid-d<sub>8</sub> and caffeic acid-13C<sub>9</sub>) before injection of 2  $\mu$ L onto a an Agilent UPLC system (Infinity 1290) equipped with a UPLC column (Acquity HSS T3, 2.1 x 50 mm, 1.8  $\mu$ m C18 in combination with a 2.1 mm x 5 mm, 1.8  $\mu$ m VanGuard precolumn; Waters Corporation, Milford, MA, USA). The UPLC system was coupled to an Agilent 6550 iFunnel Jet stream electrospray ion source Accurate-Mass QTOFMSMS (Agilent Technologies, Santa Clara, CA, USA). Mobile phases used were MilliQ water with 0.1 % formic acid (A) and 75:25 acetonitrile: 2-propanol with 0.1 % formic acid (B). The following gradient was used: 0.1 - 10 % B for 2 min, then B was increased to 99 % for 5 min and held at 99 % for 2 min; then B was decreased to 0.1 % for 0.3 min and the flow-rate was increased to 0.8 mL/min for 0.5 min; these conditions were held for 0.9 min, after which the flow-rate was reduced to 0.5 mL/min for 0.1 min before the next injection. The analysis was done in positive ion mode.

A reference interface was connected for accurate mass measurements; the reference ions purine (4  $\mu$ M) and HP-0921 (Hexakis(1H, 1H, 3H-tetrafluoropropoxy)phosphazine) (1  $\mu$ M) both purchased from Agilent Technologies (Santa Clara, CA, USA) were infused directly into the MS at a flow rate of 0.05 mL/min for internal calibration, and the monitored ions were purine m/z 121.05 and HP-0921 m/z 922.0098. The gas temperature was set to 150°C, the drying gas flow to 16 L/min and the nebulizer pressure 35 psi. The sheath gas temp was set to 350°C and the sheath gas flow to 11 L/min. The capillary voltage was set to 4000 V in positive ion mode, and to 4500 V in negative ion mode. The nozzle voltage was 300 V. The fragmentor voltage was 380 V, the skimmer 45 V and the OCT 1 RF Vpp 750 V. The collision energy was set to 0 V. The m/z range was 70 - 1700, and data were collected in centroid mode with an acquisition rate of 4 scans/s (1977 transients/spectrum).

#### LCMS Lipidomics Analysis

An aliquot (2  $\mu$ L) of the sample mixture was injected into an Agilent UPLC system (Infinity 1290) equipped with a UPLC column (Acquity CSH, 2.1 x 50 mm, 1.7 µm C18 in combination with a 2.1 mm x 5 mm, 1.7 µm VanGuard CSH precolumn; Waters Corporation, Milford, MA, USA). The UPLC system was coupled to an Agilent 6540 iFunnel Jet stream electrospray ion source QTOFMSMS (Agilent Technologies, Santa Clara, CA, USA). Mobile phases used were 60:40 ACN:water + 10 mM ammonium formate + 0.1% formic acid (A) and 89.1:10.5:0.4 IPA:ACN:water + 10 mM ammonium formate + 0.1% formic acid (B). The following gradient was used: 0.0 - 15 % B at a flow rate of 0.5 mL/min, then B was increased to 30 % during 1.2 min, then to 55% during 0.3 min and held at 55% for 3.5 min, then to 72% during 2 min, to 85% during 2.5 min and to 100% during 0.5 min where was held at 0.5 ml/min flow rate for additional 2 min after which the flow rate was increased to 5ml/min and kept for 0.5 min to wash the injection valve. After 12 min data was no longer acquired by the mass spectrometer. The column was equilibrated with 15% B for 1.5 min at a flow rate of 0.5 ml/min before the next sample was injected. The column oven temperature was held at 65°C and samples were kept in autosampler at 10°C. Analysis was performed in the positive ion mode. A reference interface was connected to facilitate accurate mass measurements; both of the reference ions (purine (4 μM) and HP-0921 (Hexakis(1H, 1H, 3H-tetrafluoropropoxy)phosphazine) (1 μM) were purchased from Agilent Technologies (Santa Clara, CA, USA)) were infused directly into the MS at a flow rate of 0.08 mL/min for internal calibration, and the monitored ions were purine m/z 121.05 and HP-0921 m/z 922.0098. The gas temperature was set to 300°C, the drying gas flow to 8 L/min and the nebulizer pressure 40 psi. The sheath gas temp was set to 350°C and the sheath gas flow to 11 L/min. The capillary voltage was set to 4000 V in positive ion mode. The nozzle voltage was 0 V. The fragmentor voltage was 100 V, the skimmer 45 V and the OCT 1 RF Vpp 750 V. The collision energy was set to 0 V. The m/z range was 70 - 1700, and data were collected in centroid mode with an acquisition rate of 4 scans/s.

## Data Processing for Lipidomics and Metabolomics Analyses

Non-processed files from GC-TOF-MS analysis were exported in NetCDF format to a MATLABbased in-house script where all data pre-treatment procedures such as baseline correction, chromatogram alignment, and peak deconvolution were performed. Metabolite identification was implemented within the script and was based on the retention index (RI) values and on MS spectra from the in-house mass spectra library established by the Swedish Metabolomics Centre (Umeå, Sweden).

Targeted feature extraction of the LCMS data was performed using the Profinder<sup>M</sup> software package, version B.06.00 (Agilent Technologies Inc., Santa Clara, CA, USA), together with a local retention time and mass spectra library containing data on 713 metabolites or 487 lipid species. Feature detection was based on the following parameters: peak height  $\geq$  300 counts; allowed ion species in positive ionization mode +H, +Na, +K, +NH4; allowed ion species in negative ionization mode -H, +HCOO; peak spacing tolerance = 0.0025-7 ppm; isotope model – common organic molecules; charge state = 1; mass tolerance = 10 ppm; retention time tolerance = 0.1 min. After peak extraction, each compound was manually checked for accurate mass and retention time agreement with appropriate standards from the library; peaks with bad characteristics (overloaded, noisy, not Gaussian etc.) were excluded from the analysis. Identification of the compounds was confirmed by comparison of their MS/MS spectra with MS/MS spectra of relevant standards. (Level 1 identification according to the Metabolomics Standards Initiative (Salek et al. 2013).

## Extraction of oxylipins

A previously reported solid phase extraction (SPE) protocol was used to isolate oxylipins from plasma samples. In summary, extraction was performed on Waters Oasis HLB cartridges (60 mg of sorbent, 30  $\mu$ m particle size). These were first washed with ethyl acetate (1 mL) and MeOH (2x2 mL), then conditioned with 5% MeOH in water (containing 0.1% acetic acid), before loading the sample. 200 microliters of each sample was spiked with 10  $\mu$ L internal standard (IS) mixture containing 50 ng/mL 12(13)-DiHOME-d<sub>4</sub> and 12(13)-EPOME-d<sub>4</sub>, 25 ng/mL 9-HODE-d<sub>4</sub>, PGE<sub>2</sub>-d<sub>4</sub> and TXB<sub>2</sub>-d<sub>4</sub>, 800 ng/mL 2-AG-d<sub>8</sub>, 40 ng/mL PGF<sub>2α</sub>-EA-d<sub>4</sub> and PGE<sub>2</sub>-EA-d<sub>4</sub>, 20 ng/mL AEA-d<sub>4</sub>, OEA-d<sub>4</sub>, SEA-d<sub>3</sub> and PEA-d<sub>4</sub> as well as 10  $\mu$ L antioxidant solution (0.2 mg/mL BHT/EDTA in methanol/water (1:1)). After applying the sample, IS, and antioxidant solution, the SPE cartridge was washed, dried under high vacuum and eluted with 2 mL methanol, 3 mL acetonitrile and 1 mL ethyl acetate into polypropylene tubes containing 6  $\mu$ L of a glycerol solution (30% in methanol). Eluates were evaporated using a MiniVac system (Farmingdale, NY, U.S.A.), reconstituted in 100  $\mu$ L of MeOH, spiked with 10  $\mu$ L recovery standard (25 ng/mL CUDA), transferred to vials and analyzed by LC-MS/MS.

### LCMS analysis of oxylipins

An LC-MS/MS instrument, the Agilent Ultra-Performance (UP)LC system (Infinity 1290) coupled to an Agilent 6490 Triple Quadrupole with an electrospray ionization source (ESI) equipped with the iFunnel Technology (Agilent Technologies, Santa Clara, CA, USA), was used with separate injections for subsequent ionization in positive (endocannabinoid) and negative (oxylipin) mode. Chromatographic separation was done on a Waters BEH C<sub>18</sub> column (2.1 mm x 150 mm, 130 Å, 1.7  $\mu$ m particle size) with an online filter, and injection volume was 10  $\mu$ L for each run. The eluents in the mobile phase consisted of (A) 0.1 % acetic acid in MilliQ water and (B) acetonitrile:isopropanol (90:10). The following gradient was employed: 0.0-3.5 min 10-35% B, 3.5-5.5 min 40% B, 5.5-7.0 min 42% B, 7.0-9.0 min 50% B, 9.0-15.0 min 65% B, 15.0-17.0 min 75% B, 17.0-18.5 min 85% B, 18.5-19.5 min 95% B, 19.5-21 min 10 % B, 21.0-25.0 min 10% B.

The ESI-MS conditions were: capillary and nozzle voltage at 4000 V and 1500 V, drying gas temperature 230 °C with a gas flow of 15 L/min, sheath gas temperature 400 °C with a gas flow of 11 L/min, the nebulizer gas flow was 35 psi, and iFunnel high and low pressure RF were set at 90 and 60 V (negative mode) and 150 and 60 V (positive mode). Dynamic multiple reaction monitoring (MRM) mode was used with fixed time windows (retention time±2 min) to profile two transitions per compound (one quantitative and one qualitative). The dynamic MRM option was performed for all compounds with optimized transitions and collision energies. The MassHunter Workstation software was used for instrument control and for manual integration of all peaks.

# Standards and calibration curve preparation

The stable isotope dilution method was used to quantify oxylipins. Two types of internal standard were used: i) for quantification purposes, deuterated IS was added before extraction, and ii) for the purpose of monitoring the loss of IS, the recovery standard CUDA was added after extraction. Eight IS were used for oxylipin quantification (12,13-DiHOME-d<sub>4</sub>, 12(13)-EpOME-d<sub>4</sub>, 9-HODE-d<sub>4</sub>, PGE<sub>2</sub>-d<sub>4</sub>, PGD<sub>2</sub>-d<sub>4</sub>, 5-HETE-d<sub>8</sub>, 20-HETE-d<sub>6</sub> and TXB<sub>2</sub>-d<sub>4</sub>). For each native compound, a suitable IS was selected based on structural similarities. Standard solutions were

prepared at 10 different levels to determine calibration curves by the least-squares linear regression model with equal weighting factor using the equation y=mx+b, where y corresponds to the response ratios (native standard peak area/internal standard peak area), m is equal to the slope of the curve, x corresponds to the on column concentration of the analyte and b is the y-interception of the calibration curve.





C)

**Fig. S1**. First locally joint loading vector (p(corr)) for the metabolites (A) and oxylipins (B) and second locally join loading vector (p(corr)) for the metabolites (C) and oxylipins (D) data sets colored according to the chemical classes (metabolites) and biochemical pathways (oxylipins).





**Fig. S2**. Pathway enrichment analysis performed for the metabolites with p(corr) values above 0.2 for first locally joint (A) and second locally joint (B) loadings between oxylipin and metabolic data sets. Metabolic pathways are arranged according to the scores from enrichment analysis (y axis) and from topology analysis (x axis). Size of the circle is connected to the value of pathway impact whereas its color to pathway's p value.

# **Reference:**

Salek, R. M., C. Steinbeck, M. R. Viant, R. Goodacre, W. B. Dunn (2013). The role of reporting standards for metabolite annotation and identification in metabolomic studies. Gigascience 2, 13 doi:10.1186/2047-217X-2-13.