CHEST Online Supplement

Alterations in Polyamine Metabolism in Patients With Lymphangioleiomyomatosis and Tuberous Sclerosis Complex 2-Deficient Cells

Yan Tang, PhD; Souheil El-Chemaly, MD; Angelo Taveira-Dasilva, MD; Hilary J. Goldberg, MD; Shefali Bagwe, MBBS; Ivan O. Rosas, MD; Joel Moss, MD, PhD; Carmen Priolo, MD, PhD; and Elizabeth P. Henske, MD

CHEST 2019; 156(6):1137-1148

 DOI: 10.1016/j.chest.2019.05.038

e-Appendix 1.

Supplementary Methods

Biological samples

Patient plasma samples were collected as previously described [11].

LAM patient-derived 621-101 were cultured at 37 degrees in 5% CO_2 in DMEM supplemented with 10% FBS, 100 mg/mL of penicillin and 100 mg/mL of streptomycin, and gentamycin sulfate (50 mg/mL). Experiments were performed on passage 24. Sixty mm plates in triplicates of 80% confluent cells were treated 24 hours with DMSO (control), rapamycin (20nM), chloroquine (5uM), and combination of rapamycin and chloroquine. Cells were washed with PBS and collected by cold 80% methanol (stored at -80 degree).

LAM lung transplant tissue was dissociated into a single cell suspension using Miltenyi Biotec human tumor dissociation kit, and immediately barcoded using 10x genomics 3' Chromium v2.0 platform. Expression library construction was performed according to manufacturer's protocol and sequenced on an illumina NextSeq500.

Metabolomics profiling

Sample Accessioning: Following receipt, samples were inventoried and immediately stored at -80°C. Each sample received was accessioned into the Metabolon LIMS system and assigned by the LIMS a unique identifier that was associated with the original source identifier only. This identifier was used to track all sample handling, tasks and results. The samples (and all derived aliquots) were tracked by the LIMS system. All portions of any sample were automatically assigned their own unique identifiers by the LIMS when a new task was created; the relationship of these samples was also tracked. All samples were maintained at -80°C until processed.

Sample Preparation: Samples were prepared using the automated MicroLab STAR® system from Hamilton Company. Several recovery standards were added prior to the first step in the extraction process for QC purposes. To remove protein, dissociate small molecules bound to protein or trapped in the precipitated protein matrix, and to recover chemically diverse metabolites, proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000) followed by centrifugation. The resulting extract was divided into five fractions: two for analysis by two separate reverse phase (RP)/UPLC-MS/MS methods with positive ion mode electrospray ionization (ESI), one for analysis by RP/UPLC-MS/MS with negative ion mode ESI and one for analysis by HILIC/UPLC-MS/MS with negative ion mode ESI. One sample was reserved for backup. Samples were placed briefly on a TurboVap® (Zymark) to remove the organic solvent. The sample extracts were stored overnight under nitrogen before preparation for analysis.

QA/QC: Several types of controls were analyzed in concert with the experimental samples: a pooled matrix sample generated by taking a small volume of each experimental sample (or alternatively, use of a pool of well-characterized human plasma) served as a technical replicate throughout the data set; extracted water samples served as process blanks; and a cocktail of QC standards that were carefully chosen not to interfere with the measurement of endogenous compounds were spiked into every analyzed sample, allowed instrument performance monitoring and aided chromatographic alignment. Instrument variability was determined by calculating the median relative standard deviation (RSD) for the standards that were added to each sample prior to injection into the mass spectrometers. Overall process variability was determined by calculating the median RSD for all endogenous metabolites (i.e. non-instrument standards) present in 100% of the pooled matrix samples. Experimental samples were randomized across the platform run with QC samples spaced evenly among the injections.

Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectroscopy (UPLC-MS/MS): All methods utilized a Waters ACQUITY ultra-performance liquid chromatography (UPLC) and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer operated at 35,000 mass resolution. The sample extract was dried then reconstituted in solvents compatible to each of the four methods. Each reconstitution solvent contained a series of standards at fixed concentrations to ensure injection and chromatographic consistency. One aliquot was analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophilic compounds. In this method, the extract was gradient eluted from a C18 column (Waters UPLC BEH C18-2.1x100 mm, 1.7 µm) using water and methanol, containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA). Another aliquot was also analyzed using acidic positive ion conditions, however it was chromatographically optimized for more hydrophobic compounds. In this method, the extract was gradient eluted from the same aforementioned C18 column using methanol, acetonitrile, water, 0.05% PFPA and 0.01% FA and was operated at an overall higher organic content. Another aliquot was analyzed using basic negative ion optimized conditions using a separate dedicated C18 column. The basic extracts were gradient eluted from the column using methanol and water, however with 6.5mM Ammonium Bicarbonate at pH 8. The fourth aliquot was analyzed via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1x150 mm, 1.7 µm) using a gradient consisting of water and acetonitrile with 10mM Ammonium Formate, pH 10.8. The MS analysis alternated between MS and data-dependent MSn scans using dynamic exclusion. The scan range varied slighted between methods but covered 70-1000 m/z. Raw data files are archived and extracted as described below.

Bioinformatics: The informatics system consisted of four major components, the Laboratory Information Management System (LIMS), the data extraction and peak-identification software, data processing tools for QC and compound identification, and a collection of information interpretation and visualization tools for use

by data analysts. The hardware and software foundations for these informatics components were the LAN backbone, and a database server running Oracle 10.2.0.1 Enterprise Edition.

LIMS: The purpose of the Metabolon LIMS system was to enable fully auditable laboratory automation through a secure, easy to use, and highly specialized system. The scope of the Metabolon LIMS system encompasses sample accessioning, sample preparation and instrumental analysis and reporting and advanced data analysis. All of the subsequent software systems are grounded in the LIMS data structures. It has been modified to leverage and interface with the in-house information extraction and data visualization systems, as well as third party instrumentation and data analysis software.

Data Extraction and Compound Identification: Raw data was extracted, peak-identified and QC processed using Metabolon's hardware and software. These systems are built on a web-service platform utilizing Microsoft's .NET technologies, which run on high-performance application servers and fiber-channel storage arrays in clusters to provide active failover and load-balancing. Compounds were identified by comparison to library entries of purified standards or recurrent unknown entities. Metabolon maintains a library based on authenticated standards that contains the retention time/index (RI), mass to charge ratio (m/z), and chromatographic data (including MS/MS spectral data) on all molecules present in the library. Furthermore, biochemical identifications are based on three criteria: retention index within a narrow RI window of the proposed identification, accurate mass match to the library +/- 10 ppm, and the MS/MS forward and reverse scores between the experimental data and authentic standards. The MS/MS scores are based on a comparison of the ions present in the experimental spectrum to the ions present in the library spectrum. While there may be similarities between these molecules based on one of these factors, the use of all three data points can be utilized to distinguish and differentiate biochemicals. More than 3300 commercially available purified standard compounds have been acquired and registered into LIMS for analysis on all platforms for determination of their analytical characteristics. Additional mass spectral entries have been created for structurally unnamed biochemicals, which have been identified by virtue of their recurrent nature (both chromatographic and mass spectral).

Metabolite Quantification and Data Normalization: Peaks were quantified using area-under-the-curve. To avoid batch effect, all profilings were run within same day. Thus, no normalization is necessary. Biochemical data have been normalized to total protein as determined by Bradford assay to account for differences in metabolite levels due to differences in the amount of material present in each sample.

Lipidomics Platform

TrueMass® Lipomic Panel: Lipids were extracted in the presence of authentic internal standards using chloroform:methanol (2:1 v/v). For the separation of neutral lipid classes [FFA, TAG, DAG, CE], a solvent system consisting of petroleum ether/diethyl ether/acetic acid (80:20:1) was employed. Individual phospholipid classes within each extract [PC, PE] were separated using the Agilent Technologies 1100 Series

LC. Each lipid class was transesterified in 1% sulfuric acid in methanol in a sealed vial under a nitrogen atmosphere at 100°C for 45 minutes. The resulting fatty acid methyl esters were extracted from the mixture with hexane containing 0.05% butylated hydroxytoluene and prepared for GC by sealing the hexane extracts under nitrogen. Fatty acid methyl esters were separated and quantified by capillary GC (Agilent Technologies 6890 Series GC) equipped with a 30 m DB 88 capillary column (Agilent Technologies) and a flame ionization detector.

TrueMass® Ceramides Panel: Deuterium-labeled internal standards were added to the samples and the mixture was solubilized in methanol followed by a crash extraction. A bilayer was formed with the addition of KCl in water, and the organic layer was removed and concentrated under nitrogen. The extract was spun, filtered, and split into 2 injections – one for ceramides and one for sphingosines. The extract was injected onto an Agilent C8 column connected to an Agilent 1290 Infinity LC and ABI 4000 QTRAP. The analytes were ionized via positive electrospray and the mass spectrometer was operated in the tandem MS mode. The absolute concentration of each sphingolipid was determined by comparing the peak to that of the relevant internal standard.

Statistical Analysis

Metabolomics Data processing. Missing data was filtered out. Values present across above 70% of samples were retained. All the missing and zero values were replaced with the minimum value across all samples assuming to be the detection limit. The assumption of this approach is that most missing values are caused by low abundance metabolites (i.e. below the detection limit). Data was log2 transformed and normalized by Supervised Normalization of Microarray (SNM) [15]. Individual effect was removed by setting patients as adjustable variable. Treatment conditions of baseline, during treatment, and post-treatment were adjusted by setting as biological variable.

MetaboAnalyst Analysis. Significance Analysis of Microarray (SAM) was used for biomarker identification distinguishing treatment effect. SAM assigns a significance score to each variable based on its change relative to the standard deviation of repeated measurements. For a variable with scores greater than an adjustable threshold, its relative difference is compared to the distribution estimated by random permutations of the class labels. For each threshold, a certain proportion of the variables in the permutation set is found to be significant by chance. The proportion is used to calculate the FDR [17].

Empirical Bayesian Analysis of Microarray (EBAM) was used for biomarker identification distinguishing treatment effect. EBAM is an empirical Bayesian method based on moderated t-statistics. EBAM uses a two-group mixture model for null and significant features [17].

Partial Least Squares - Discriminant Analysis (PLS-DA) was used for biomarker identification distinguishing treatment effect. PLS is a supervised method that uses multivariate regression techniques to extract via linear combination of original variables the information that can predict the class membership [17].

Self-organizing map (SOM). Hexagonal fingerprint was generated using log2 transformed values with the R supraHex package [18].

Differential analysis. R limma [16] was used for differential analysis.

LAM Lung primary culture. Lung tissue was dissociated into single cell suspension by gentleMACS™ Octo Dissociator with Heaters within one hour. Dissociated cells were cultured in DMEM supplemented with 10% FBS for 2 weeks and passaged into two 10cm plates until 80% confluent. Cell cultures were treated with 20nM rapamycin dissolved with DMSO or DMSO (1ul/10 culture medium) as control for 24 hours. Cells were trypsinized and cell viability (95%) was assessed by trypan blue staining and Invitrogen countess cell counter.

Droplet-based scRNA-Seq. Immediately after trypsinization, cells were barcoded with 10x genomics ChromiumTM Single Cell 3' Library & beads kit V2, aiming for approximately 4,000 cells captured per library. scRNA-Seq libraries were constructed parallel in all steps following 10x genomics Single Cell 3' Reagent Kits v2 instruction. Libraries were sequenced on Illumine Nextseq 500.

scRNA-Seq data processing. The CellRanger software (http://10xgenomics.com/) was used for sample demultiplexing, barcode processing and gene expression counting. Paired reads were mapped to the hg38 human reference genome. Mean reads measured per cell was around 60,000 for both DMSO control group and rapamycin treatment group. For DMSO control group, 3,005 cells were captured, and median genes measured per cell was around 3,500 genes; for rapamycin treatment group, 2,700 cells were captured, and median genes measured per cell was around 3,600 genes. Data normalization and TPMs were computed using R Seurat [12]. Log2 transformed values were used for analysis. Low quality cells were filtered out by threshold of 15% mitochondria RNA content. Cells expressing less than 200 genes were filtered out, and genes expressing in less than 3 cells were filtered out. Cell types were annotated using R SingleR package against Blueprint encode database [21] and then manually checked by expression of well-studied fibroblasts marker genes. Annotation of LAM-activated fibroblasts (LAFs) were based on the expression of a panel of marker genes for activated fibroblasts, including FAP, PDPN, MMP2, MMP11, PDGFRA, PDGFRL, CTSK. The mean value of averaged expressions was set as cutoff threshold. Cells with score higher than the threshold were annotated as LAFs.

e-Figure 1. Variation analysis of metabolomic profiles.

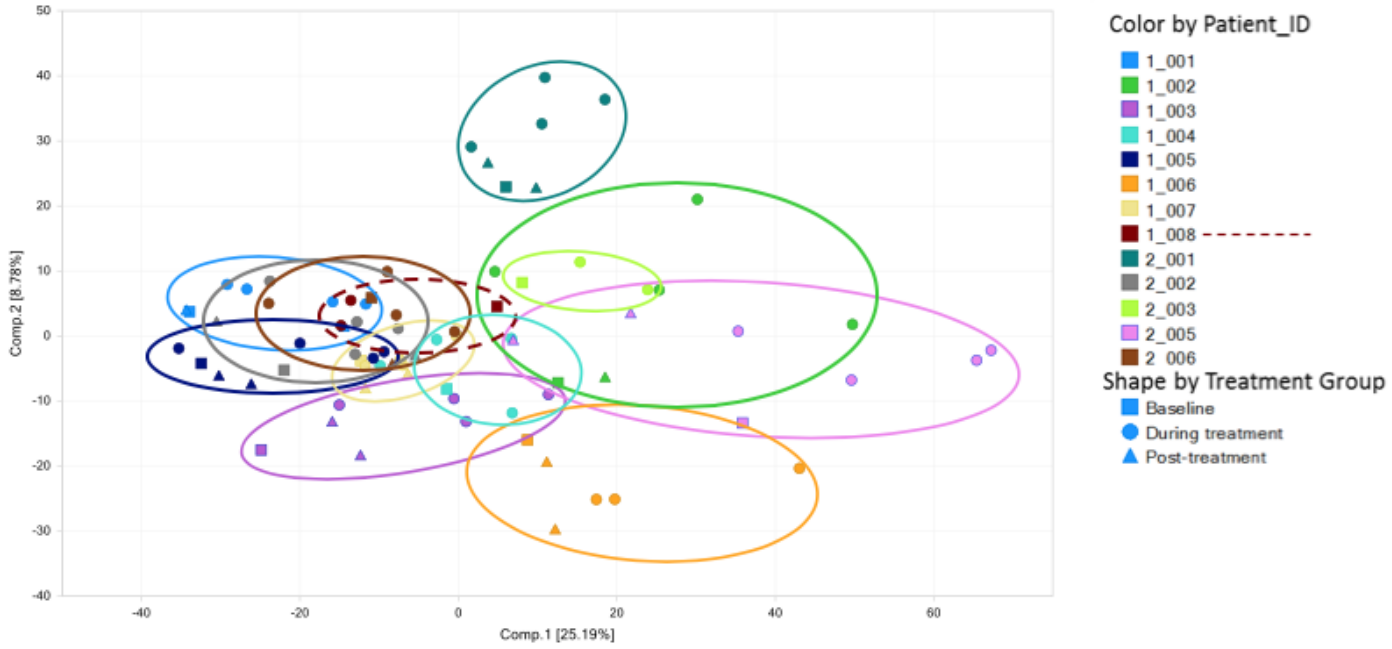
(A) Before data normalization, the primary driver of clustering in first two principle components was individual subject. Each color represents an individual patient.

(B-C) After data normalization by Supervised Normalization of Microarray (SNM), treatment effects were uncovered.

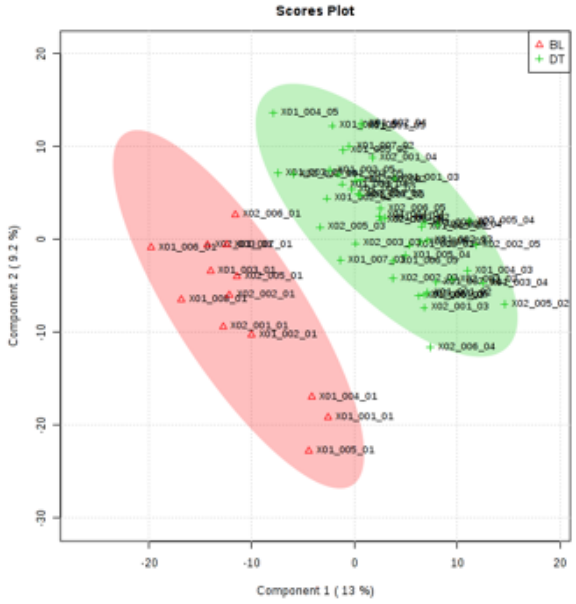
(B) Partial Least Square-Discriminant Analysis (PLS-DA) of metabolomic profiles comparing During Treatment (DT) time points (green) with baseline (BL) (red).

(C) Partial Least Square-Discriminant Analysis (PLS-DA) of metabolomic profiles comparing Post Treatment (PT) time points (green) with During Treatment (DT) time points (red).

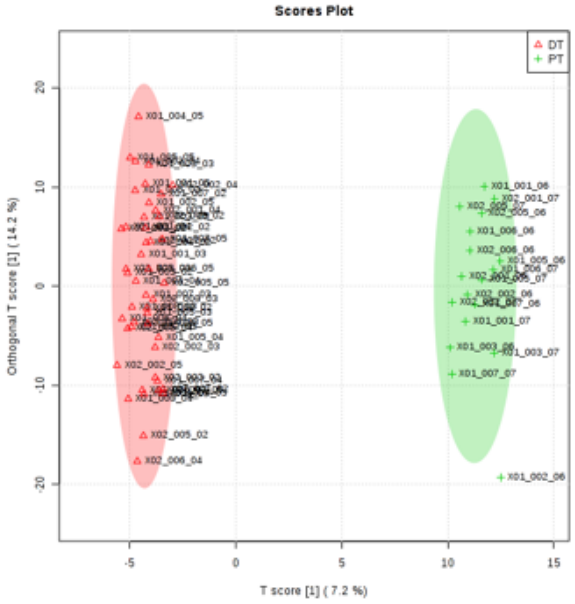
A



B



C

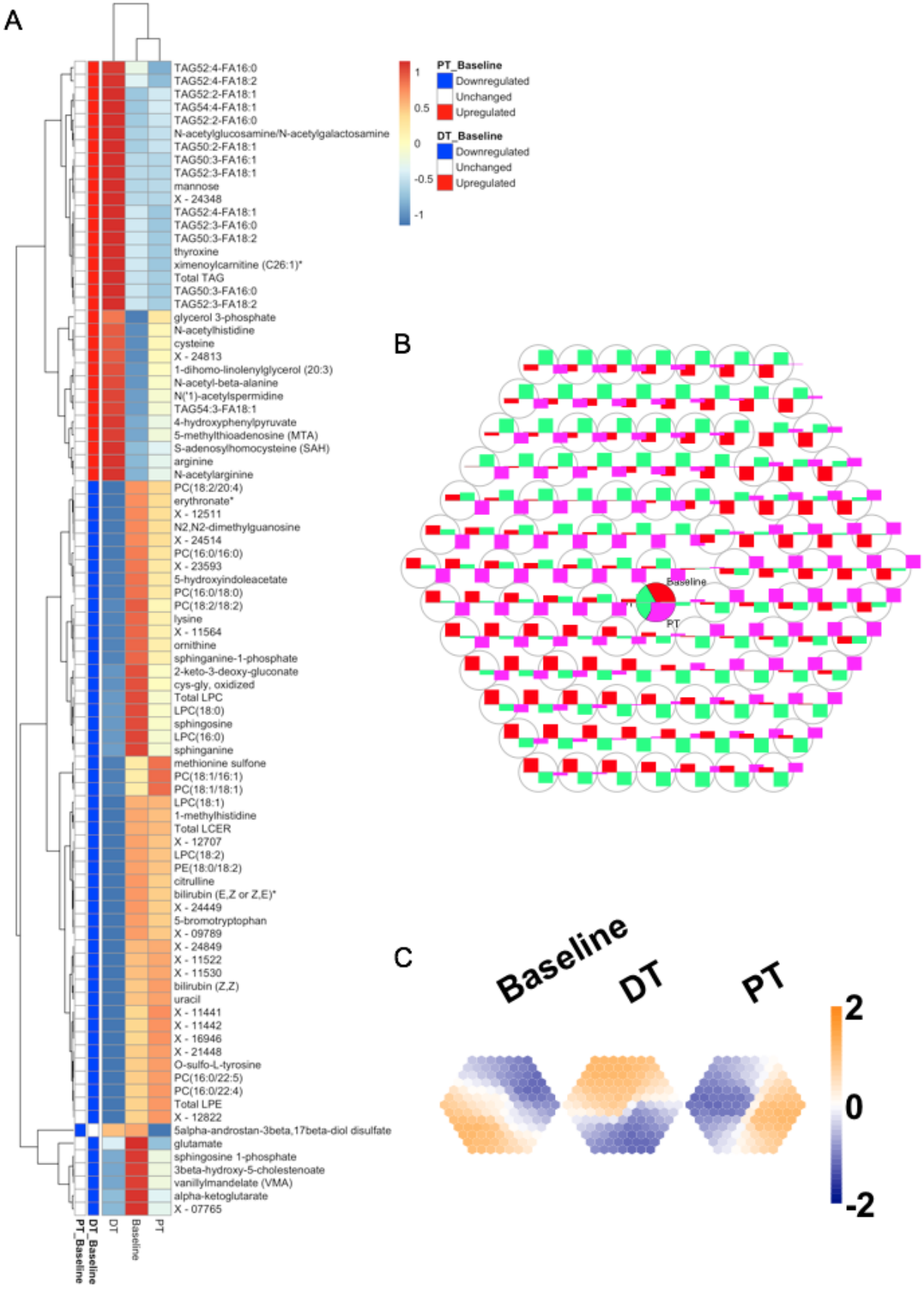


e-Figure 2. Plasma metabolomic profiling reveals metabolic alterations by combinatorial treatment of sirolimus and hydroxychloroquine.

(A) Heatmap of significantly upregulated or downregulated metabolites during treatment (DT) and post treatment (PT). The two bars at the far left represent significant downregulation (blue) or upregulation (red) during treatment (DT) or post treatment (PT).

(B) Metabolite Fingerprinting by Self-Organizing Map (SOM). Each cluster (hexagon) contains metabolic features that share similar pattern across all subjects at baseline (red), during treatment (DT, green), and post treatment (PT, pink). Data are scaled. Bars above the horizontal line represents upregulation, and bars below the line represents downregulation.

(C) SOMs across treatment conditions shows altered metabolomic patterns during treatment (DT) and post treatment (PT) compared to baseline. Color represents upregulation (up to 2, orange) or downregulation (down to -2, blue) of metabolites in each cluster at 3 treatment stages.



c

(See separate files for e-Tables 1-4.)

e-Table 5. List of metabolites identified by EBAM as biomarkers separating treatment from baseline.

Compounds	z.value	posterior	local.fdr
arginine	6.5559	0.99997	2.70E-05
N1-methyladenosine	-6.2038	1	4.42E-06
O-sulfo-L-tyrosine	-5.9264	0.99999	1.34E-05
pseudouridine	-5.6861	0.99997	3.50E-05
N2,N2-dimethylguanosine	-5.5768	0.99995	5.38E-05
5-methylthioadenosine MTA	5.168	0.99932	0.00067658
sphinganine-1-phosphate	-4.928	0.99937	0.00063434
methionine sulfone	-4.9168	0.99934	0.00066068
2-keto-3-deoxy-gluconate	-4.8389	0.99913	0.0008748
Total HCER	-4.8267	0.99909	0.00091373
dimethylarginine SDMA ADMA	-4.6936	0.99854	0.0014614
erythronate	-4.6816	0.99848	0.0015236
bilirubin E,Z or Z,E	-4.6512	0.99831	0.0016926
sphingosine 1-phosphate	-4.4788	0.99697	0.0030321
citrulline	-4.4617	0.99679	0.0032091
5-bromotryptophan	-4.4453	0.99661	0.0033862
sphinganine	-4.3683	0.99565	0.0043492
aspartate	-4.357	0.99549	0.0045104
ergothioneine	-4.2583	0.99384	0.0061626
glycerol 3-phosphate	4.2523	0.99488	0.0051192
7-alpha-hydroxy-3-oxo-4-cholestenoate 7-Hoca	-4.2391	0.99346	0.0065409
CE203	4.2061	0.99435	0.0056499
bilirubin Z,Z	-4.1974	0.99256	0.0074371
sphingosine	-4.1915	0.99243	0.0075733
cys-gly, oxidized	-4.1112	0.99036	0.0096443
7-methylguanine	-4.0928	0.98982	0.010184
3beta-hydroxy-5-cholestenoate	-4.0799	0.98942	0.010579
deoxycarnitine	-3.9741	0.98564	0.014358
PC180/226	3.9268	0.98984	0.010161
CE205	3.9268	0.98984	0.010162
S-adenosylhomocysteine SAH	3.921	0.98972	0.010284
PC180/203	3.8665	0.98849	0.011511
thyroxine	3.8568	0.98826	0.011744
CE226	3.8392	0.98782	0.012176
alpha-ketoglutarate	-3.7566	0.97414	0.025862

SM201	3.7066	0.98404	0.015959
SM240	3.7033	0.98393	0.016068
CE204	3.6709	0.98285	0.017154
N-acetyl-beta-alanine	3.6585	0.98241	0.017589
isoleucine	3.6352	0.98157	0.018435
ornithine	-3.6048	0.96227	0.037726
SM241	3.5901	0.97982	0.020176
CE183	3.531	0.9773	0.022698
N-acetylalanine	-3.4962	0.95144	0.048562
PC180/204	3.4945	0.9756	0.024399
biliverdin	-3.4353	0.94443	0.055569
methylmalonate MMA	-3.4237	0.94302	0.056983
SM221	3.4184	0.97166	0.028341
CE182	3.3839	0.96969	0.030312
Total CE	3.3797	0.96944	0.030564

e-Table 6. Details of significant features identified by SAM. Arginine and MTA were among top features.

Compounds	d.value	stdev	rawp	q.value
arginine	2.6129	0.23852	0	0
N2,N2-dimethylguanosine	-2.3505	0.26223	0	0
N1-methyladenosine	-2.3475	0.21911	0	0
O-sulfo-L-tyrosine	-2.2951	0.2275	0	0
5-methylthioadenosine MTA	2.2113	0.26919	0	0
sphinganine-1-phosphate	-2.0979	0.26683	0	0
Total HCER	-2.0578	0.26751	0	0
2-keto-3-deoxy-gluconate	-2.0368	0.26164	0	0
bilirubin E,Z or Z,E	-2.0142	0.27492	0	0
pseudouridine	-2.0112	0.19699	0	0
sphingosine 1-phosphate	-1.995	0.28911	0	0
erythronate	-1.9875	0.26554	0	0
sphinganine	-1.953	0.29105	1.45E-05	0.00046062
citrulline	-1.9278	0.27384	1.45E-05	0.00046062
5-bromotryptophan	-1.9189	0.27339	4.34E-05	0.0012091
methionine sulfone	-1.9176	0.23015	4.34E-05	0.0012091
sphingosine	-1.8863	0.29454	5.79E-05	0.001433
aspartate	-1.8634	0.26897	8.68E-05	0.0020364
dimethylarginine SDMA ADMA	-1.8475	0.23364	0.0001013	0.0021496
3beta-hydroxy-5-cholestenoate	-1.8459	0.2974	0.0001013	0.0021496
7-methylguanine	-1.8326	0.29185	0.00013025	0.0024183
7-alpha-hydroxy-3-oxo-4-cholestenoate 7-Hoca	-1.8305	0.27356	0.00013025	0.0024183
ergothioneine	-1.8305	0.27138	0.00013025	0.0024183
cys-gly, oxidized	-1.8009	0.28058	0.00015919	0.0027283
deoxycarnitine	-1.7839	0.29317	0.00015919	0.0027283
bilirubin Z,Z	-1.7738	0.26343	0.00017366	0.0027637
alpha-ketoglutarate	-1.7015	0.29802	0.00030391	0.0046697
ornithine	-1.6268	0.29604	0.00052098	0.0064487
biliverdin	-1.5554	0.29781	0.00098408	0.010198

(See separate file for e-Table 7.)