Methods and Protocols

Animals

All mouse work was performed in accordance with the Institutional Animal Care and Use Committees (IACUC) and relevant guidelines of Boston Children’s Hospital. Embryonic day (E) 14.5 embryos from time-pregnant CD-1 (ICR) dams were purchased from Charles River Laboratories (CRL) for the baseline conditions metabolomics database and maternal kynurenine injection time-course experiments. E12.5 embryos from timed-pregnant C57BL/6 dams were purchased from CRL for maternal immune activation metabolomics experiments. For timed pregnancies, females were checked for the presence of plugs, and the date of the plug was noted as embryonic day 0.5 (E0.5). All animals were housed under 12hr/12hr day night cycle with access to standard chow and water ad libitum.

Embryonic and maternal body fluid and tissue collection and analysis

Embryonic CSF (eCSF) was collected by inserting a glass capillary into the cisterna magna and processed as described (Zappaterra et al, 2013). Each CSF sample was pooled across litters, purity was verified visually with a dissection microscope, and then frozen on dry ice. Embryonic blood was collected by glass capillary following a neck artery nick and pooled across litters. Maternal blood was collected either by tail vein nick for non-terminal timepoints (e.g. prior to a 48h end-point collection) puncture for terminal timepoints (e.g. at 3h end-point collection or at 48h end-point collection). Serum samples were analyzed following coagulation, centrifugation, and then immediately put on dry ice. Embryonic and maternal livers were dissected and immediately transferred to dry ice.

20

Maternal immune activation and validation

On E12.5, pregnant dams received a single dose (20 mg/kg, i.p.) of polyinosinic-polycytidylic acid (polyI:C, Sigma Aldrich) or sterile saline as vehicle control. To confirm maternal immune activation, maternal blood was collected at 3h following maternal polyI:C injection by either tail- nick or intracardiac puncture. Maternal blood was also collected 48h following maternal polyI:C injection by intracardiac puncture. After maternal blood samples coagulated and centrifuged, 3h and 48h MIA serum samples were analyzed by LEGENDplex Mouse Inflammation Panel (13- plex with V-bottom Plate) FACS-ELISA (BioLegend). Validation of MIA was performed by measurements of cytokine levels in maternal serum using FACS-ELISA. Results reported in EV3B are from mice that are the source of the data of experiment 1 harvested at 48h that is reported in Figure 3G, I. Results reported in EV3D are from mice that are the source of the data of experiment 2 harvested at 48h that is reported in Figure 3H, J. Statistical analyses were performed using unpaired t-tests through the GraphPad Prism 9.2.0 and Prism 10 software.

Sample preparation for LC-MS analysis of polar metabolites from CSF, plasma, and liver from embryonic and adult mice
We used fresh frozen tissues or processed (see method above) CSF and plasma. Per condition, 5- 10 μl embryonic CSF, 5 μl plasma or a grain-of-rice sized sample of maternal and embryonic livers were extracted. Tissues were homogenized with a handheld homogenizer (Sigma, Z359971) and briefly sonicated. 300μl of extraction solvent (80% LC/MS-grade methanol, 20% 25 mM Ammonium Acetate and 2.5 mM Na-Ascorbate prepared in LC/MS water and supplemented with isotopically labeled internal standards (17 amino acids and isotopically labeled reduced glutathione, Cambridge Isotope Laboratories, MSK-A2-1.2 and CNLM-6245- 10)) was used per sample. After centrifugation for 10 min at maximum speed on a benchtop centrifuge (Eppendorf) the cleared supernatant was dried using a nitrogen dryer (ThermoFisher Scientific, TS-18826) and reconstituted in 17 μl (eCSF) or 25ul (plasma, tissues) “QReSS” water (supplemented with QReSS, Cambridge Isotope Laboratories, MSK-QRESS-KIT) by brief vortexing. Extracted metabolites were spun again and the cleared supernatant was transferred to LC-MS micro vials. A small amount of each sample was pooled and serially diluted 3 and 10- fold for quality controls (pool-QC) throughout the run of each batch. About 1⁄2 of eCSF (7μl),

21

1/25th of plasma (1μl), and 1/25th of tissue (1μl) were injected onto a chromatographic column and subsequently analyzed using LCMS.

Chromatographic conditions for LC-MS

ZIC-pHILIC chromatography was used for polar metabolites: an appropriate volume of reconstituted sample was injected into a ZIC-pHILIC 150 × 2.1 mm (5 μm particle size) column (EMD Millipore) operated on a VanquishTM Flex UHPLC Systems (Thermo Fisher Scientific, San Jose, CA). Chromatographic separation was achieved under the following conditions: buffer A was acetonitrile; buffer B was 20 mM ammonium carbonate, 0.1% ammonium hydroxide. Gradient conditions were: linear gradient from 20% to 80% B; 20–20.5 min: from 80% to 20% B; 20.5–28 min: hold at 20% B. The column oven and autosampler tray were held at 25 °C and 4 °C, respectively. Chromatographic performance was quality controlled using a mixture of unlabeled standard amino acids and a mixture of chemically diverse compounds (Cambridge Isotope Laboratories, MSK-A2-US-1.2 and MSK-QRESS-US-KIT) with 1μl of each injected before or after every run on our HILIC method.

MS data acquisition conditions for targeted or untargeted analysis of polar metabolites

MS data acquisition was performed using a QExactive benchtop orbitrap mass spectrometer equipped with an Ion Max source and a HESI II probe (Thermo Fisher Scientific, San Jose, CA) and was performed in positive and negative ionization mode in a range of m/z= 70–1000, with the resolution set at 70,000, the AGC target at 1x106, and the maximum injection time (Max IT) at 20 msec. Optimal HESI conditions were determined utilizing a mix of chemical standards: Sheath gas flow rate: 35; Aug gas flow rate: 8; Sweet gas flow rate: 1; Spray voltage: 3.5kV (pos), 2.8kV (neg); Capillary temperature: 320oC; S-lens RF: 50; Aux gas heater temperature: 350 oC. For untargeted analysis, additional Top8 method (see below) was employed in either positive or negative mode with ddMS settings as: the resolution set at 17,500, the AGC target at 1x105, the maximum IT at 50 msec, isolation window at 1.0 m/z, and stepped NCE at 20,40 and 80.

As our software/hardware did not allow for the widely employed AquireX assisted MS2 inclusion/exclusion list data collection, we performed repeated data collections with the following logic: At the end of each experimental batch spectral data was collected in both

22

positive and negative mode on additional injections of the pool-QC. We collected spectral data in a data-dependent acquisition mode (DDA) such that for every MS1-level scan the top 8 most abundant ions were picked for fragmentation (Top8-level experiment), an exclusion window was set to 6 seconds, such that ions picked for fragmentation would be put on an exclusion list for 6 seconds before they could be picked for fragmentation again. This allowed for the collection of about 300-500 individual ion spectral scans per sample injection. Then, MS1 and MS2 level data were quickly analyzed using CompoundDiscoverer and detected features were used to create an inclusion and exclusion list (per polarity) that ensured maximal inclusion of interesting features (features with signals above mock level and with no MS2 spectra collected) and maximal exclusion of non-interesting features (features present in mock samples for which MS2 spectra was already collected). This strategy was then performed iteratively up to 4 times. We noted an incremental increase in the number of features with associated spectra generated with this strategy, however we note that this strategy did not perform as well as reported from an AquireX type of experiment.

Generation of targeted metabolomics libraries (in-house MS1 and MS2-level databases)

Two separate MS1 and MS2-level libraries were generated. A set of standards premixed in 7 individual pools were run on our LC-MS HILIC method. Information on this pool set has been previously published (Petrova et al, 2021; Sullivan et al, 2019). Briefly, we first collected MS1 level information (retention time and peak shape and quality were noted, see Table EV1 see also Dataset EV3). To generate the MS2-level database, each pool set was injected twice (for positive and negative mode), and specific inclusion lists were preset for a PRM-type of experiment. This inclusion list was based on the already determined retention times and was focused on compounds with peaks of minimum acceptable quality. Thus, high quality MS2 data were collected per individual metabolite and MS2 spectra were averaged out from three different HCD cell energies: 20, 40, 80 NCE. We named this database “HILIC\_all” (Dataset EV1). The second library set was based on the MSMLS library of compounds (MSMLS, IROA Technologies’), which we named “MSMLS\_HILIC”. Compounds were solubilized and pooled based on manufacturer’s recommendations. Each pool was run as an individual injection and MS1 information (retention time and peak shape and quality) was noted. (Table EV2 and see also Dataset EV4). Then, spectral MS2 information was collected at both ionization modes as above

23

with the exception that individual spectra were not averaged but collected at the three different HDC energies independently (Dataset EV2). All spectra were manually inspected and compared to known and well characterized spectral databases (METLIN, MONA and mzCloudTM). Most of the spectra included in our library were thus verified. Where no match was found but the spectrum was of high quality, we included it in the library and added a note in our metadata file (Dataset EV3 (HILIC\_all) and EV4 (MSMLS\_HILIC). Few metabolites were also run from individual wells, to try and obtain better signal-to-noise ratios. Overall, our database included: HILIC-all MS1 – 261 compounds, MS2 – 144 compounds; MSMLS\_HILIC MS1 – 236 compounds; MS2 – 307 compounds. The overlap between the two MS1 databases was about 100 compounds.

Targeted metabolomics data analysis

Relative quantification of polar metabolites was performed with TraceFinder 5.1 (Thermo Fisher Scientific, Waltham, MA, USA) using a 5 ppm mass tolerance and referencing an in-house library of chemical standards (HILIC\_all and MSMLS\_HLIC). We routinely queried about 257 compounds (37 internal standards and 220 metabolites). Pooled samples and fractional dilutions were prepared as quality controls and injected at the beginning and end of each run. In addition, pooled samples were interspersed throughout the run to control for technical drift in signal quality as well as to serve to assess the coefficient of variability (CV) for each metabolite. Data from TraceFinder were further consolidated and normalized with an in-house R script: (https://github.com/FrozenGas/KanarekLabTraceFinderRScripts/blob/main/MS\_data\_script\_v2.4 \_20221018.R). Briefly, this script performs normalization and quality control steps: 1) extracts and combines the peak areas from TraceFinder output .csvs; 2) calculates and normalizes to an averaged factor from all mean-centered chromatographic peak areas of isotopically labeled amino acid and QReSS internal standards within each sample; 3) filters out low-quality metabolites based on user entered cut-offs calculated from pool reinjections and pool dilutions; 4) calculates and normalizes for biological material amounts based on the total integrated peak area values of high-confidence metabolites. In this study, the linear correlation between the dilution factor and the peak area cut-offs are set to RSQ>0.95 and the coefficient of variation (CV) < 30%. Finally, data were log transformed and Pareto scaled within the MetaboAnalyst- based statistical analysis platform (Xia et al, 2015) to generate PCA, PLSDA, volcano plots, and

24

heatmaps. Individual metabolite bar plots and statistics were calculated in Excel and GraphPad Prism. Finally, for statistical analysis within MetaboAnalyst missing values were filled using LoDs (1/5 of the minimum positive value of each variable) while when plotting individual metabolites (using GraphPad Prism) missing values were omitted from analysis.

Untargeted metabolomics data analysis using CompoundDiscoverer (mock-filtering mode)

Relative quantification for untargeted polar metabolomics was performed with Compound discoverer 3.3. A general workflow was built to best suit our polar metabolomics LC-MS method (details on each parameter are given in Dataset EV5). Positive and negative modes were analyzed separately. Both MS1 and MS2-level in-house databases were used (HILIC\_all and MSMLS\_HLIC). For the analysis of MIA samples, additionally MS1 and MS2-level CD-1-DBs were used (in positive and negative modes, see above). Filtering steps were performed based on peak noise levels, ppm error, formula annotation, and the relative abundance of the integrated peak area in true sample compered to mock samples, where features with +>3-fold higher in samples were retained. Normalization steps were performed based on a “scaling factor”: a per sample normalizer. This scaling factor was determined per experiment or sample comparison: each sample was analyzed by targeted metabolomics first, and the scaling factor corresponded to the biological material normalizer from total integrated peak area of mean-centered values from high-confidence metabolites (see above). This normalizer rarely differed more than 2-fold within a sample type. When experiments from different experimental days were compared (such as CD- 1 to C57BL/6 comparison, as in Figure 2), untargeted analysis was performed on the combined samples and data were independently curated. This ensured adequate chromatographic retention time correction and alignment and thus proper feature alignment, identification, and annotation. Within our HILIC chromatography, we rarely observe chromatographic shift larger that 1 minute and retention times drifts larger than 40 seconds. We further observed that larger threshold for retention time shift led to mis-annotation for our in-house library compound databases (for example betaine vs valine or alanine vs beta-alanine). Therefore, these parameters were set as limits for retention time correction and matching to internal databases. Post CompoundDiscoverer analysis included merging of negative and positive mode data and statistical analysis using the MetaboAnalyst platform. As compound annotation at levels higher than Level 1 include a significant level of uncertainty, we decided not to merge features based on

25

annotation name but rather to retain information from both positive and negative mode analyzed data (thus multiple compounds are annotation as either “Metabolite\_name\_[M-H]” or “Metabolite\_name” for negative or positive mode respectively, for example Fig 2D). Finally, data were log transformed and Pareto scaled within the MetaboAnalyst-based statistical analysis platform (Xia et al, 2015) to generate PCA, PLSDA, volcano plots, and heatmaps. Individual metabolite bar plots and statistics were calculated in Excel and GraphPad Prism. Pathway enrichment was similarly performed within MetaboAnalyst following the above data transformation steps.

Untargeted metabolomics data analysis using ClusterFinder and IROA credentialing (IROA mode)
All steps for sample preparation for IROA-assisted metabolomics were as described above except for the final resuspension step, which was performed using the IROA provided Internal Standard (IS) (95% 13C-labeled, IROA® TruQuant Yeast Extract Semi-targeted QC Workflow, IROA TechnologiesTM). Relative quantification based on IROA for untargeted polar metabolomics was performed with ClusterFinder 4.2.23. We initially determined that 7ul of CSF reconstituted in 20μl IS led to optimal signal intensity correspondence between labeled and unlabeled isotopologues (see Figure S1, step 1). Default parameters were set for each search (targeted or untargeted) except for the minimal signal intensity threshold which was set lower, at 10.000. Additionally, we integrated our targeted metabolomics MS1-level databases into ClusterFinder and thus performed a secondary level of annotation within the software. Retention time correction and peak integration was performed within the software and was not further curated.

Generation of untargeted metabolomics CD-1 library (CD-1-DB MS1 and MS2-level database)

Embryonic CSF from E14.5 embryos was collected as described above. 7 mice were used and CSF from all embryos was pooled per mouse. For each replicate, 7μl were used for IROA- assisted analysis and 7μl for mock-filtering analysis. As IROA analysis did not yield satisfactory annotation we only generated eCSF-specific metabolite databases from mock-filtering strategy. Within CompoundDiscoverer, we followed the steps outlined above. We extensively curated each feature. Features were given identifying tags that would help organize certainty of

26

identification and annotation (Level 1-4). All features with matches to our in-house database were confirmed after manual inspection and any double annotations removed (as well as renamed and specifically tagged) if retention times were not ideal match. Both MS1 and MS2 level databases were generated within CompoundDiscoverer on curated data. After export of MS2 library to mzVault, the CD-1-eCSF database was further manually curated, where low quality spectra were removed. Of note – most features are represented by a single collected spectrum.

UNTARGETED CERTAINTY TAGS:

CD-1-DB dataset 03.25

|  |  |  |
| --- | --- | --- |
| A | MSlist match in both HILIC\_all and MSMLS\_HILIC MS1 databases | Corresponds to Level 1 identification certainty |
| B | mzVault match (multiple and single) in both HILIC\_all and MSMLS\_HILIC MS2 databases | Corresponds to Level 1 identification certainty |
| C | NA | NA |
| D | mzVault match (multiple and single) in MSMLS\_HILIC MS2 database | Corresponds to Level 1 identification certainty |
| E | mzCloud match (greater than or equal to 65 confidence) | Corresponds to Level 2 identification certainty |
| F | Peak quality filter | Bad peaks removed manually |

*All other untargeted datasets:*

|  |  |  |
| --- | --- | --- |
| A | MSlist match in both HILIC\_all and MSMLS\_HILIC MS1 databases | Corresponds to Level 1 identification certainty |
| B | mzVault match (multiple and single) in both HILIC\_all and MSMLS\_HILIC MS2 databases | Corresponds to Level 1 identification certainty |
| C | mzVault match (multiple and single) in CD-1-DB MS2 | Corresponds to Level 2 identification certainty (with MS spectrum) |
| D | mzCloud match (greater than or equal to 65 confidence) | Corresponds to Level 2 identification certainty |
| E | up or down regulated |  |
| F | MSlist match in CD-1-DB MS1 | Corresponds to Level 3 identification certainty (no MS spectrum) |