# Methods

Sample collection

Cerebrospinal fluid was collected from a total of 25 patients and 5 healthy controls. Patients presented mutations in the genes *CDKL5* (n=2), *GRIN* (n=7), *MECP2* (n=12), and *STXBP1* (n=4). The median age at the time of collection was 6 years, ranging from infancy to 18 years. The majority of patients were female (n=19) due to the large proportion of Rett syndrome patients that were used in this study. One patient with a *GRIN* mutation was excluded from the study due to the severity of her symptoms and her early dead.

|  |  |
| --- | --- |
| Characteristic |  |
| Age, years, median (range) | 6 (0-18) |
| Sex, female, n (%) | 19 (76) |
| Intellectual disability, n (%) | 19 (76) |
| Encephalopathy, n (%) | 17 (68) |
| Epilepsy, n (%) | 18 (72) |
| Autism, n (%) | 3 (12) |

Metabolomics analysis

Samples were sent to the Centre for Omic Sciences (COS) Joint Unit of the Universitat Rovira i Virgili-Eurecat, where a targeted metabolomics analysis was conducted by liquid chromatography with tandem mass spectrometry (LC-MS-MS) for tryptophan related metabolites and acylcarnitines. Gas chromatography–mass spectrometry (GC-MS) was performed for organic acids and sugar metabolites related to energy metabolism and the TCA cycle. A total of 82 metabolites were detected, and the results were presented as micromolar (μM) concentrations.

LC-MS/MS methodology

CSF (50 μL) samples were diluted in methanol (200 μL, 100%) and the set of labelled internal standards (Supplementary information). Following centrifugation (10 min, 15000 rpm, 4ºC), the supernatant was dried in a SpeedVac concentrator and reconstituted in methanol (50 μL, 100%).

The equipment consisted on an UHPLC 1290 Infinity II Series coupled to a QqQ/MS 6490 Series (Agilent Technologies) and a Kinetex 2.6 μm Polar C18, 100 Å, 150 x 2.1 mm (Phenomenex) analytical column. Chromatographic separation was performed in mode negative electrospray ionization (ESI) at 20ºC with an injection volume of 1 μL. Mobile phase A was formic acid (0.1%) and mobile phase B was methanol:formic acid (10:1, v/v).

GC-MS methodology

CSF (50 μL) samples were diluted in water:methanol (200 μL, 8:2, v/v) and the set of labelled internal standards (Supplementary information). Following centrifugation (5 min, 15000 rpm, 4ºC), the supernatants were dried in a SpeedVac concentrator at 45ºC and reconstituted in methoxyamine (30 μL, 100%). Samples were incubated at 37ºC for 90 min. and then silylated with 45 μL of MSTFA + 1 % at room temperature for 60 min.

The separation was performed on an GC-QTOF 7200 and a HP5-MS UI capillary column (30 m x 250 μm I.D., 0.25 μm film thickness), both from Agilent Technologies. Helium (>99.999%) was used as the carrier gas with a constant flow 1.1 mL/min. Initial oven temperature was set at 60ºC, then increased by 10ºC/min. to 320ºC and held constant for 10 min. Samples were injected in split mode 1:20 at injection temperature 250ºC. Compounds were detected through MS in electron ionization (70 eV) and full-scan monitoring mode (m/z 50–600) mode using the spectra library Fiehn-pct-2013. Ion source temperature was 250ºC and quadrupole temperature was 200ºC. The acquisition rate was 5 spectra/s.

Statistical analysis

The goal of the analysis was to compare the metabolic profiles of patients to those of healthy controls. The diseases presented by the patients were categorized based on their glutamatergic component. All statistical tests were conducted in R version 4.3.0.

The metabolite concentrations were log2-transformed in order to account for the expected high variability and noise found in metabolomics data. Multivariate statistical methods were applied on the transformed data in order to account for the small sample size and the inherent multicollinearity of the data. Unsupervised principal component analysis (PCA) was performed in order to obtain an overall view of the variation between samples and their separation into clusters. Ward hierarchical clustering with Euclidian distance combined with a heatmap (pheatmap version 1.0.12) was used to compare the log-transformed concentrations of all the analyzed metabolites in patients and controls. Orthogonal partial least-squares discriminant analysis (OPLS-DA) from the ropls package (version 1.32.0) was used as a supervised model to identify metabolites responsible for group separation. Models were evaluated by goodness of fit (R2Y), goodness of prediction (Q2Y), and the root mean square error of estimation (RMSEE). 7-fold cross-validation was performed to mitigate the risk of overfitting caused by the high dimensionality of the data. Metabolites were selected based on their variable importance in projection (VIP) scores (VIP > 1). Permutation testing (n=1000) was used to determine the statistical significance of the results (p<0.05).

The results of the classification models were corroborated by performing Wilcoxon Mann–Whitney tests comparing each metabolite concentration in patients and in controls. The non-parametric test was chosen due to the small number of samples. Correction for the false discovery rate (FDR) was applied in order to account for the issue of multiple testing (p<0.05).

Pathway analysis was performed for metabolites with p<0.05 or VIP score > 1 using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and the KEGGREST package (version 1.40.0). Pathways were identified as important when two or more of their metabolite components were found to be altered between the disease groups and the healthy controls. The results were plotted after excluding pathways that are ubiquitous or unrelated to synaptic metabolism. Special attention was given to the tryptophan metabolism pathway, as well as to the amino acid substrates of the ABC transporters.

LO PIDEN LOS DE EURECAT

Acknowledgements

We thank Núria Canela, Antoni del Pino Rius, Elisabet Foguet, Salvador Fernández Arroyo, and Iris Samar

ra Mestre from the Metabolomics facility of the Centre for Omic Sciences (COS) Joint Unit of the Universitat Rovira i Virgili-Eurecat, for their contribution to mass spectrometry analysis.