Metabolic network changes that are strongly associated with Dementia with Lewy Bodies as determined through signed distance and partial correlation analysis

Miroslava Čuperlović-Culf[1-3*], Ali Yilmaz[4,8], Sumeyya Akyol[8], Sangeetha Vishweswaraiah[8], David Stewart[1], Anuradha Surendra[1], Xiaojian Shao[1], Irina Alecu[2,3], Thao Nguyen-Tran[2,3], Bernadette McGuinness[5], Peter Passmore[5], Patrick G. Kehoe[6], Michael E. Maddens[4,8], Brian D. Green[7], Steffany A.L. Bennett[2,3*], Stewart F. Graham[4,8*]

[1] National Research Council of Canada, Digital Technologies Research Centre, Ottawa, Canada
[2] Ottawa Institute of Systems Biology, Ottawa, Ontario, Canada
[3] Department of Biochemistry, Microbiology, sand Immunology, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada
[4] Oakland University-William Beaumont School of Medicine, Rochester, MI 48309, USA
[5] Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, UK
[6] Dementia Research Group, Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
[7] Institute for Global Food Security, School of Biological Sciences, Queen's University Belfast, UK
[8] Beaumont Research Institute, Royal Oak, MI 48073, USA

Correspondance:
Miroslava Cuperlovic-Culf e-mail:miroslava.cuperlovic-culf@nrc-cnrc.gc.ca;
Stewart F. Graham, e-mail: stewart.graham@beaumont.org and
Steffany A.L. Bennett: steffanyann.bennett@uottawa.ca

ABSTRACT

INTRODUCTION

Determining significant metabolic changes in Dementia with Lewy Bodies (DLB), a complex and multifactorial neurodegenerative disease, requires, in addition to the analysis of concentration changes, a deep understanding of functional modifications in the context of metabolic networks.

METHODS

Brain metabolomics data from DLB patients and healthy controls was explored using novel correlation analysis approaches to identify metabolites with the largest changes in their network in the disease state.

RESULTS

Novel clustering and correlation network analysis shows major change in the metabolic network in DLB brain relative to age and sex-matched controls with the largest interaction network alterations for fructose, propylene-glycol, pantothenate and *O*-acetylcarnitine, although there were no statistically significant changes in their concentrations.

DISCUSSION

Network and correlation analyses indicate major changes in the purine degradation pathway, propanoate and β-alanine metabolism as well as an increased role of fructose and reduced significance of glucose in the brains of DLB patients.

1. Introduction

With 10-15% of dementia patients ultimately diagnosed with an underlying aetiology of Dementia with Lewy bodies (DLB), it is considered the most prevalent cause of dementia after Alzheimer's disease (AD). Neurodegeneration due to DLB is age-associated and is clinically characterized by cognitive fluctuations, extrapyramidal motor symptoms and early visual hallucinations[1]. Prodromal, early stage DLB shows similar symptoms to other dementias making precise, early diagnosis of DLB particularly difficult with individuals often misdiagnosed with PD, multiple system atrophy (MSA) or AD[2].

Objective diagnostic tests and therapies other than those for symptoms associated with DLB are currently non-existent. Accurate diagnosis is further complicated by major differences in symptoms and disease presentations amongst patients and the expressed variability over time[3]. While some new information on the etiopathogenesis of DLB holds promise, progress towards the development of efficacious disease-modifying therapies is curtailed due to a lack of understanding of the early underlying pathological mechanisms. As mentioned, the limited available treatment options primarily focus on the reduction of individual symptoms which often leads to numerous side effects unfortunately with only minimal benefits[4]. Discovery of molecular biomarkers which can accurately diagnose DLB in its infancy would help stratify patients correctly and eventually lead to the development of novel therapies and prevention modalities which could be implemented earlier, when disease-modifying strategies are likely to be effective.

The major role of metabolites and the metabolome in neurodegenerative diseases are being increasingly recognized. For example, several recent reports highlight metabolic alterations in brains of patients who died from DLB. Notably, changes in mitochondria and energy metabolism, purine metabolism and protein synthesis observed in the frontal cortex combined with an innate inflammatory response have been described[5,6]. In PD, another Lewy body progressive neurodegenerative disorder characterised primarily by motor symptoms, genome-wide association studies underscore lipid-associated metabolic pathways and in particular their interaction with α-synuclein, in relation to the pathogenies of the disease (reviewed recently[8,9]). Further, several specific metabolite species have been delineated for their significance in disease development and progression such as ceramides which have been reported to be important to PD development due to their roles as signaling molecules, as essential elements of cell membranes and as regulators of lipid metabolism[10]. Observed changes in metabolic profiles in dementia patients are likely to be a combination of both environmental and genetics factors.

Combined high-throughput analyses of metabolites and lipids by nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) provides more comprehensive coverage of the metabolome and as such, a greater insight into the relevance of the observed metabolite and metabolic network changes. Diversity of the metabolome, extremely large dynamic range, as well as the need to exploit multiple platforms in order to measure a larger number of compounds presents the analysis of metabolite interaction networks with unique challenges. Tools for network analysis based on graph theory have been previously applied for the analysis of metabolic networks with some early examples exploring the network properties such as degree distribution[11] indicating connectedness for each feature as well as community detection[12] differentiating networks based on feature proximity based on network connections. In metabolic network design, recent studies also report utilization of extensive knowledge of metabolism and resources such as complete known metabolic network - RECON3[13] for the development of more accurate as well as directed and flux-dependent graphs[14]. However, when analysing the metabolome and lipidome, the issues of minimally described reaction networks and sparse coverage remain major hurdles.

Recently a novel approach for correlation analysis, called distance correlation, has been introduced to the analysis of omics-based data. Distance correlation was developed by Székely et al.[15] , as a measure of dependence between variables including linear and non-linear correlations as well as correlation between vectors of different lengths. Additionally, applications of distance correlation in genomics have shown that although distance correlation calculation provides only positive values, it is possible to use sign measure from Pearson's correlation for the determination of signed distance correlation measure[16] . In addition to

correlation network development, methods from clustering can provide alternative approaches for interaction or dependence network determination. Fuzzy c-means clustering has been previously used in metabolomics[17], and its ability to assign features, e.g. genes or metabolites, to multiple groups is invaluable in biological studies, with many features involved in different processes[18]. This approach can also provide valuable information about the relationships between features through their co-clustering or classification dependence amongst groups.

Herein, we detail the novel application of signed distance correlation for the determination of functional feature changes augmented by use of fuzzy clustering and partial correlation analysis for the investigation of major metabolic changes that are associated with DLB pathogenesis.

2. Methods

2.1 Sample characteristics

Data used in this study was previously described in great detail[7]. Briefly, human brain samples, harvested from the neocortex (Brodmann region 7), have been collected from pathologically confirmed patients with DLB (n=15) and age-, and sex- matched controls with no known neurological disease (n=30). Tissues were obtained from the Brains for Dementia Research Group, Institute of Clinical Neurosciences, School of Clinical Sciences, University of Bristol, Bristol, UK as part of the study approved by the Beaumont Health System's Human Investigation Committee (HIC No.: 2018-387) and following all approved guidelines. Further details about the patients and samples characteristics are provided in the original publication[7].

2.2 Experimental procedures

2.2.1 NMR analysis

Samples were analyzed as previously described by our group[19-21]. In brief, samples previously stored at −80 °C were lyophilized and milled to a fine powder under liquid nitrogen minimizing heat production. 50 mg samples were extracted in 50% methanol/water (1 g/mL) in a sterile 2 mL Eppendorf tube. Samples were mixed for 20 min and sonicated for 20 min at 4 °C. Macromolecules and larger proteins were removed via centrifugation at 13,000×g at 4 °C for 30 min. Supernatants were collected, dried under vacuum using a Savant DNA SpeedVac (Thermo Scientific, Waltham, MA USA), and reconstituted in 285 µL of 50 mM potassium phosphate buffer (pH 7.0), 30 µL of sodium 2,2-dimethyl-2- silapentane-5-sulfonate (DSS), and 35 µL of D2O (Ravanbakhsh et al., 2015). 200 µL of the reconstituted sample was transferred to a 3 mm Bruker NMR tube for analysis. All samples were housed at 4°C in a thermostatically controlled SampleJet autosampler (Bruker-Biospin,USA) and heated to room temperature over 3 min prior to analysis by NMR. All 1D ^{1}H NMR data were recorded at 300 (±0.5) K on a Bruker Ascend HD 600 MHz spectrometer (Bruker-Biospin, Billerica, MA, USA) coupled with a 5 mm TCI cryoprobe in a randomized fashion. For each sample, 256 transients were collected as 64 k data points with a spectral width of 11.964 Hz (20 ppm), using a pulse sequence called CPP WaterSupp (Bruker pulse program: pusenoesypr1d) developed by Mercier et al. (Mercier et al., 2011) and an inter-pulse delay of 9.65 s. The data collection protocol included a 3 min equilibration period, fast 3D shimming using the z-axis profile of the ^{2}H NMR solvent signal, receiver gain adjustment, and acquisition. The free induction decay signal was zero filled to 128 k points prior to Fourier transformation, and 0.1 Hz of line broadening was applied. The singlet produced by the DSS methyl groups was used as an internal standard for chemical shift referencing (set to 0 ppm, concentration 500 µM) and for quantification. All ^{1}H NMR spectra were processed and analyzed using the Chenomx NMR Suite Professional Software package version 8.1 (Chenomx Inc, Edmonton, AB). Further details regarding sample preparation and analytical measurements and analysis have been provided previously[7].

2.2.2 Targeted mass spectrometry analysis

Metabolite quantification in postmortem brain tissue was completed as previously described by our group using the commercially available AbsoluteIDQ p180 Kit (Biocrates Life Sciences AG)[19-21]. In brief, 10 mg of lyophilized and powdered brain tissue was extracted using 300 µL of extraction solvent (85% ethanol and 15% phosphate-buffered saline solution). The samples were ultrasonicated for 10 min on ice and vortexed for 1 min. Proteins and other impurities were removed by centrifugation at 13,000×g for 15 min at 4 °C. The supernatant was collected and 10 µL was used for analysis. Data were acquired using an Acquity I-class (Waters, USA) coupled with a Xevo TQ-S mass spectrometer (Waters, USA). Metabolite concentrations were calculated using the MetIDQ software (Biocrates Life Sciences AG)

2.3 Data Analysis

Prior to analysis, all features were z-score normalized across all samples if not stated otherwise. Imputation of missing values was performed for samples and features that had less than 30% coverage using the KNN imputation method with N=10 and Euclidean distance for similarity assessment.

2.3.1 Correlation network

2.3.1.1 Distance correlation network

Correlation was calculated on z-score normalized features using Pearson, Spearman and Distance correlation analysis. Matlab routine *corr* was used for Pearson and Spearman correlation. Distance correlation was calculated following Szekely and Rizzo[15] as defined in Eq. 1:

$$dCor(X,Y) = \frac{dCov(X,Y)}{\sqrt{dVar(X)dVar(Y)}} \quad (1)$$

While Pearson correlation uses covariance between values as: $Cor(X,Y) = \frac{Cov(X,Y)}{\sqrt{\sigma_X \sigma_Y}}$ with covariance calculated as: $(X,Y) = \sum_{i=1}^{n}(x_i - \underline{x})(y_i - \underline{y})$; distance correlation depends on the distance covariance determined as: $dCov(X,Y)^2 = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} A_{j,k}B_{j,k}$ with A and B as simple linear functions of the pairwise distances between elements in samples X and Y. A and B are doubly centered distance matrices for variables X and Y respectively calculated from the pairwise distance between elements in each sample set with $A_{j,k} = a_{j,k} - \underline{a}_{j.} - \underline{a}_{k.} + \underline{a}_{..}$ where $a_{j,k} = \sqrt{(x_j - x_k)^2}$ and $\underline{a}_{j.}$ and $\underline{a}_{k.}$ respectively the *j*-row and *k*-column mean and $\underline{a}_{..}$ the overall mean of *A*. *B* includes equivalent measures to *A* for variable *Y*. Distance correlation calculation was written in-house under Matlab using *pdist2* to calculate distances between features with Euclidean metric as well as a Python application. Correlation was calculated between all features separately for sample groups. Function corr provides the p-value for the correlation measure. For distance correlation we calculated p-value using Student's t cumulative distribution function (*tcdf* function in Matlab). Distance correlation values with p>0.01 were set to zero keeping only the most statistically significant correlations. Sign of distance correlation was equated to the sign of Pearson correlation calculation as shown by[16] and implemented in-house as: $dCor(X,Y) = sign\left(\frac{Cov(X,Y)}{\sqrt{\sigma_X \sigma_Y}}\right) * \frac{dCov(X,Y)}{\sqrt{dVar(X)dVar(Y)}}$. Calculations we performed using Matlab in-house routing. Signed distance correlation calculation is also provided as a Web application at http://complimet.ca/SiDCo with RShiny front-end for a Python implementation of the method which includes z-score normalization and signed distance correlation calculation for calculation of correlation between each feature and all other features combined (correlation with the network) and pairwise correlation between features.

2.3.1.2 Partial correlation

Partial correlation was calculated using recursive method from Spearman correlation where correlation between two variables *i* and *j* with exclusion of correlations with variable *k* is calculated as: $\rho_{ij/k} = (\rho_{ij} -$

$\rho_{ik}\rho_{jk})/(\sqrt{1-\rho_{ik}^2}\sqrt{1-\rho_{jk}^2})$. Partial correlation analysis was performed for distance correlation analysis determined significant features. Calculations were performed using Matlab 2021a (Mathworks Inc).

2.3.1.3 Determination of functional feature changes

Feature (metabolite) differences between control and DLB groups based on the network results are obtained using linear regression analysis between the two groups for each metabolite comparing values of edges, i.e. pairwise correlations, to all other metabolites between two groups. The level of change is reported as a slope of linear regression. The major changes are obtained as the largest divergence from the slope of 1 obtained for identical trends in correlations with a metabolite. This analysis was used to determine major behavioural differences between patient groups.

2.3.2 Machine learning

2.3.2.1 Fuzzy C-means clustering network

In addition to the Principal Component Analysis (PCA) and hierarchical clustering (HCL) we have used Fuzzy c-means (FCM) clustering[17,22]. FCM allows each feature to belong to more than one group by providing a degree of membership, "belonging", to each cluster by maximizing proximity between similar features and distance between dissimilar features. FCM is based on the minimization of the objective function:$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^m \|x_i - c_j\|^2$ where $m \in (1,\infty)$ is the "fuzzyfication" factor, u_{ij}^m is the membership degree for feature x_i to the cluster *j* with $c_j$defining the cluster center. FCM clustering assigns objects to groups with features belonging to the same clusters showing more similarity to each other than to features in other clusters. Higher membership value indicates stronger belonging to the cluster with membership value of 1 ultimately indicating that feature is only associated with the single cluster. Calculations were performed using Matlab 2021a (Mathworks Inc).

2.3.2.2 Feature selection

Selection of the most significant features between different groups of samples F-test (using fsrftest running under Matlab) as well as machine learning method Relieff[23] (Lionelle 2005) and LASSO method for regularization of linear models[24]. Correlation results are presented using circular plots using routine circularGraph [https://github.com/paul-kassebaum-mathworks/circularGraph] with some in-house modifications. Enrichment analysis for metabolites was performed with MetaboAnalyst 5.0[25].

3. Results

3.1 General sample characteristics

Combined NMR and MS metabolomics was used to accurately measure the concentration of 218 metabolites in brain samples from DLB patients and age and sex-matched healthy controls. Sample information is summarized in Figure 1A. Preliminary Principal Component Analysis (PCA) and Hierarchical clustering (HCL) of log-corrected and z-score normalized brain metabolomics data shows some limited grouping of samples by diagnosis in PC2 but not a clear unsupervised separation by sample type (Figure 1B). Additional PCA of the two cohorts shows small correlation between PC1 and age in the control group and PC2 and sex in the DLB group (Figure 1B). Finally, hierarchical clustering (Figure 1B) performed on z-score normalized data and using Ward linkage provide clustering by chemical groups but, similarly to PCA, only minor grouping by sample types.

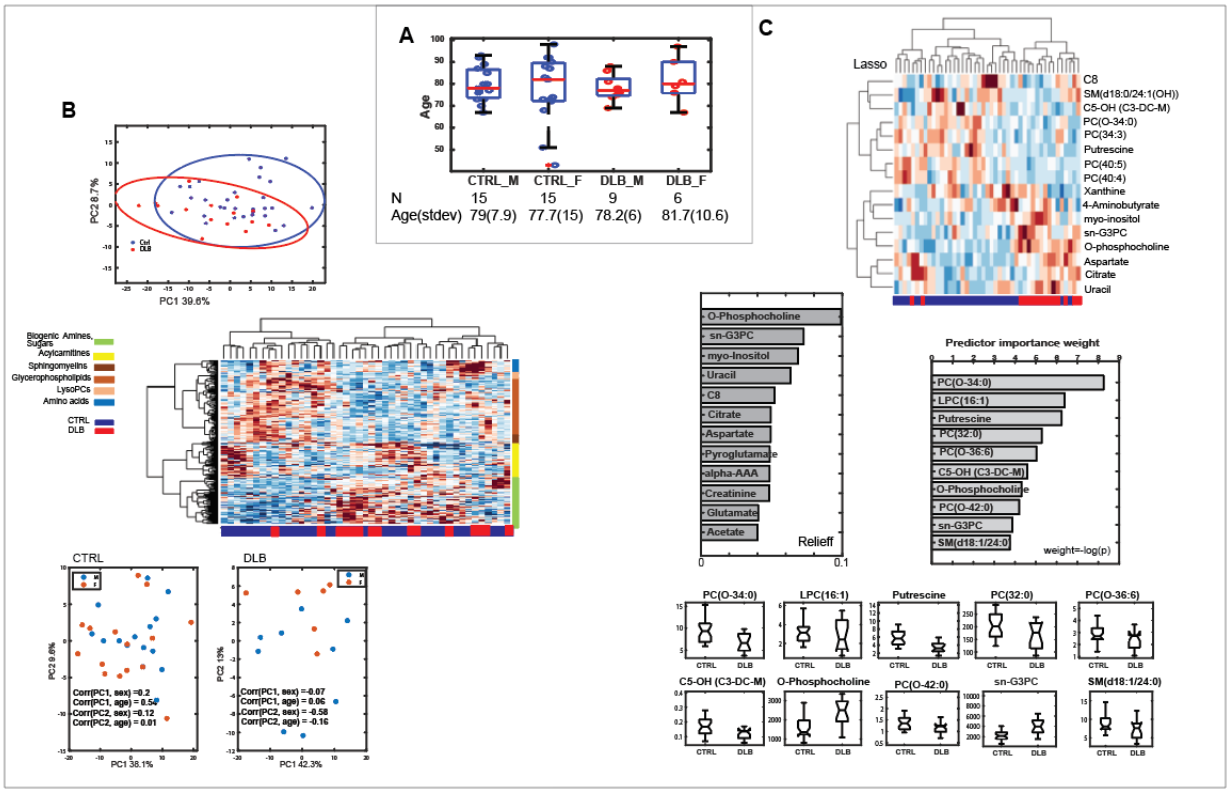


Figure 1. Unsupervised analysis does not show significant separation of control and DLB brain samples. A. Overview of age and sex representation of samples in control (CTRL) and DLB cohorts; B. Principal Component Analysis (PCA) of log transformed and z-score normalized features. Ellipses indicate 95% confidence regions; Hierarchical clustering (HCL) of samples and features following z-score normalization of features, indicated are general metabolic groups as well as sample types; PCA of two cohorts separately and Pearson correlation between PC1 and PC2 and sex and age; C. Metabolites showing the most significant differences between DLB and CTRL remaining after removal of irrelevant features using LASSO; selected for their group classification power by Relieff and determined to individually show the most significant difference between groups using F-test analysis; features with weight value over 4 (corresponding to $p<0.02$) are selected as significant and in LASSO features with weight over zero remain in selection. An improvement in sample classification with HCL is provided by LASSO selected features. Nomenclature for all metabolites is provided in Supplementary Table 2.

3.2 Feature selection and clustering

The analysis presented in Figure 1C indicates that overall differences in metabolite concentrations between DLB and control brain samples do not provide unsupervised sample separation. Specific metabolic differences between the two groups thus require further supervised analysis and feature selection as well as the analysis of differences in the behavior between metabolites.

Selection of the most significantly changed metabolite concentrations are performed using the F-statistics test and LASSO. Metabolites with statistically significant weight are shown in Figure 1C and the group of metabolites selected by LASSO to provide the most information for classification is included in Figure 1C.

The enrichment analysis by MetaboAnalyst 5.0[25] reveals that the most significantly different features between DLB and controls are part of phospholipid biosynthesis, phosphatidylcholine biosynthesis, spermidine and spermine biosynthesis, retinol metabolism, sphingolipid metabolism and methionine metabolism. Several metabolites that have been selected by both F-test and LASSO are o-phosphocholine, *sn*-glycero-3-phosphocholine (*sn*-G3PC), putrescine, PC(O-34:0) and hydroxyvalerylcarnitine (C5-OH (C3-

DC-M)). Observed concentration changes can be a result of either change in production or in utilization and network analysis of these features, as well as analysis of major network changes, can provide further information about the possible sources of observed differences.

Further metabolite grouping has been explored using fuzzy c-means clustering (FCM) (Figure 2) separately in the control and DLB groups. Although in both groups, features primarily cluster within their chemical types, in agreement with the overall clustering with HCL (Figure 1B), there are several notable differences. Overview of FCM membership using PCA visualization (Supplementary Figure 1A) shows a change in proximity, co-clustering between several groups with clearer separation of acyl carnitines, organic acids and sugars as well as amino acids in DLB, while glycerophosphocholines (PCs) and lysophosphocholines (LPCs) as well as sphingomyelins (SMs) are more crisply clustered in the control group further provided with lists of metabolites across five clusters (Supplementary Figure 1B and C).

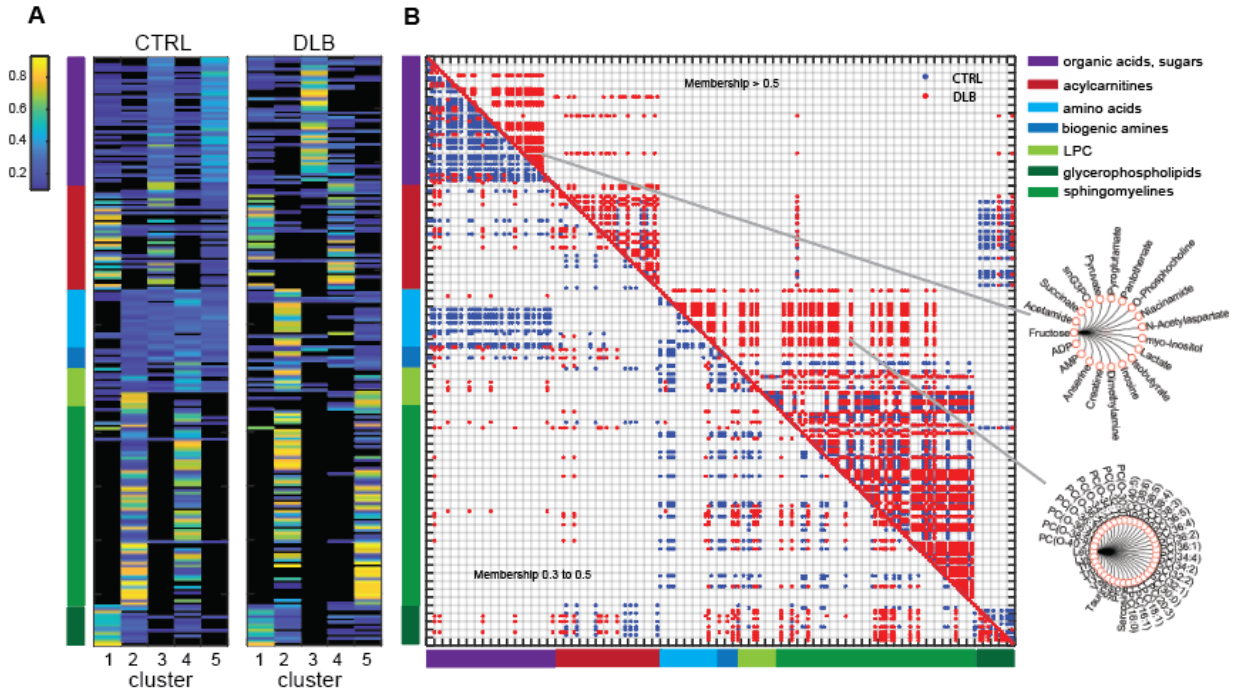


Figure 2. Fuzzy c-means clustering of metabolites in Control and DLB cohorts. FCM is calculated using fuzzyfication factor m=1.5 with 5 clusters following z-score normalization. A. Membership values over 0.1 for two cohorts in five clusters. B. FCM cluster network created by linking features through co-clustering. Shown are feature links in CTRL (blue) and DLB (red) for membership values over or equal to 0.5 (upper right triangle) and values between 0.3 and 0.5 membership (lower right triangle). Major changed in clustering of several features is indicated with strong clustering of fructose with number of metabolites as well as, amino acids with PCs and LPCs only observed in DLB group.

Investigation of individual groups shows a major shift in clustering of SM with strong association with carnitines in the control group and with PC and LPC in DLB. Contrastingly, fructose clusters with PC and LPC in control samples while in DLB fructose is strongly clustered with metabolites from galactose metabolism (Glucose, UDP-glucose, UDP-galactose, myo-inositol, D-fructose, ADP) (Figure 2).

3.3 Signed Distance correlation network analysis and explainable feature selection

Herein we detail a data-driven approach to determine changes in the metabolic profiles and networks in DLB brain. We employed distance correlation[15] to determine proximal as well as distant interactions with both non-linear and linear correlations. Combining distance correlation with Pearson's correlation (see Materials and Methods) provides signed distance correlation values indicating both strength and directionality of correlations. Overall, metabolic networks with edges at significance level under 0.01 and absolute value of correlation over 0.6 (selected as value of the largest number of connections) are shown in the spy plot. Negative correlations show major shifts in the DLB group and are presented in detail in circo plots (Figure 3).

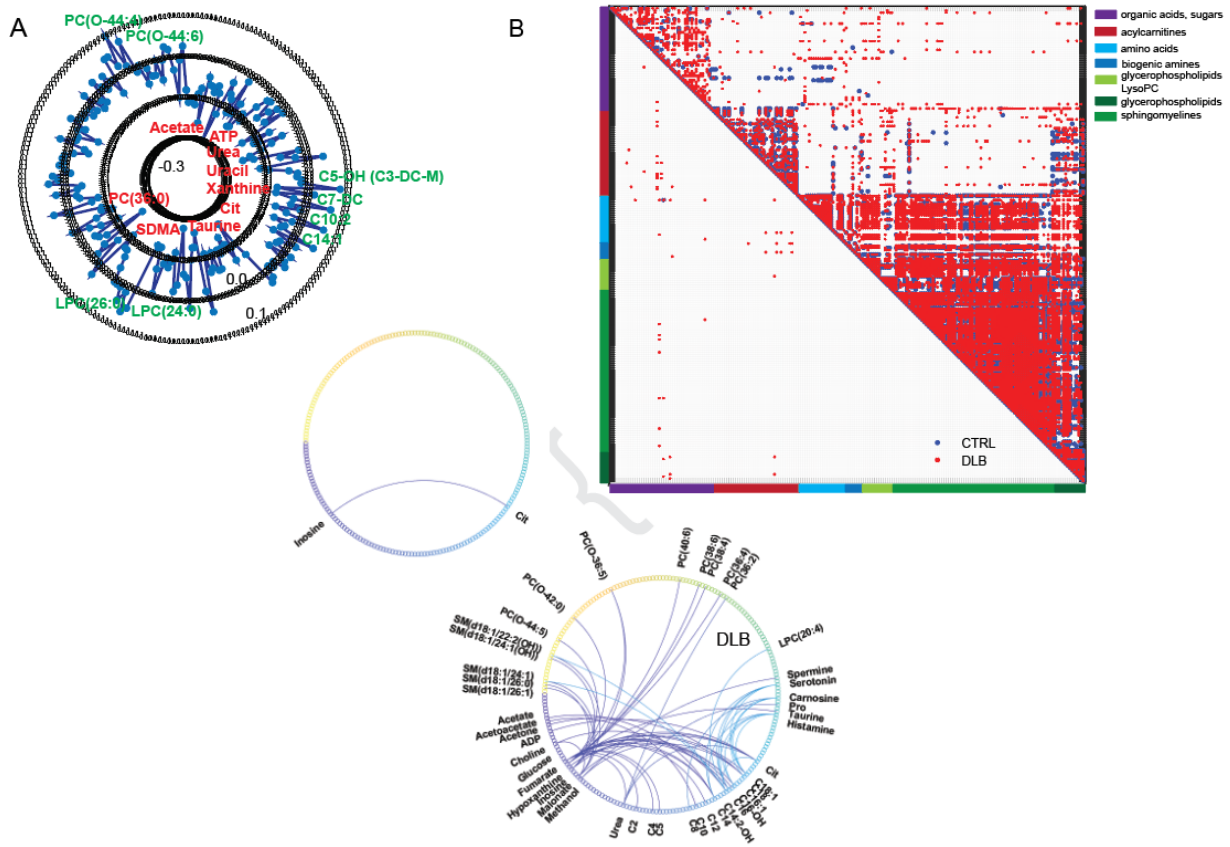


Figure 3. A. Distance correlation between each feature and all the other features is calculated for control and DLB groups. Shown is difference between values in control and DLB with metabolites with the largest differences indicated where metabolites in green have higher value in control samples. B – Pairwise, signed distance correlation. Signed distance correlation in control (blue dots and A) and DLB (red dots and B) with negative correlation shown as circo plots as well (A and B). We present correlations with an absolute value over 0.6 and $p<0.01$. The Spy plot shows positive (upper right triangle) and negative correlations (lower left triangle). Indicated are groups of metabolites in the spy plot as well as specific metabolites in the circo plots. C.

Both positive and negative correlations show significant changes in DLB with a general increase in the number of significant correlations. This is particularly striking amongst the observed negative correlations between metabolites and lipids including negative correlations between glucose and several phospholipids, SM, spermine and acetylcarnitines, in addition to urea, proline and taurine in brains from DLB patients. As distance correlation allows calculation of correlation between vectors of different length,

the correlations between each metabolite and all other metabolites combined was also calculated in both control and DLB groups, indicating correlation between a specific metabolite and the network. Differences between these values is shown in Figure 3with metabolites with the largest change indicated. According to this analysis, the biggest change in the correlation with overall metabolic network is observed for ATP, SDMA and PC(O-C44:4) and hydroxyvalerylcarnitine (C5-OH (C3-DC-M)). All values for both groups are provided in Supplementary Table 1. In this analysis, increase in DLB indicates increased absolute correlation between metabolite and the complete network.

Metabolites with the most significantly altered interaction network in the DLB group are shown in Figure 4, where Figure 4A displays the major network difference obtained from the linear regression analysis of metabolites 1-to-all correlations in DLB vs. control and Figure 4B is the opposite - control vs. DLB (see Materials and Methods). Fructose shows the largest change in its pairwise metabolic correlation network between DLB and control. Metabolic networks for the top three metabolites in both regression plots, including fructose, propylene glycol, pantothenate and *O*-acetylcarnitine are shown in Figure 4C-E. It is important to stress that none of these metabolites shows statistically significant difference in concentrations between control and DLB groups; however, their correlation changes, indicating that different pathway involvement is significant between the two groups.

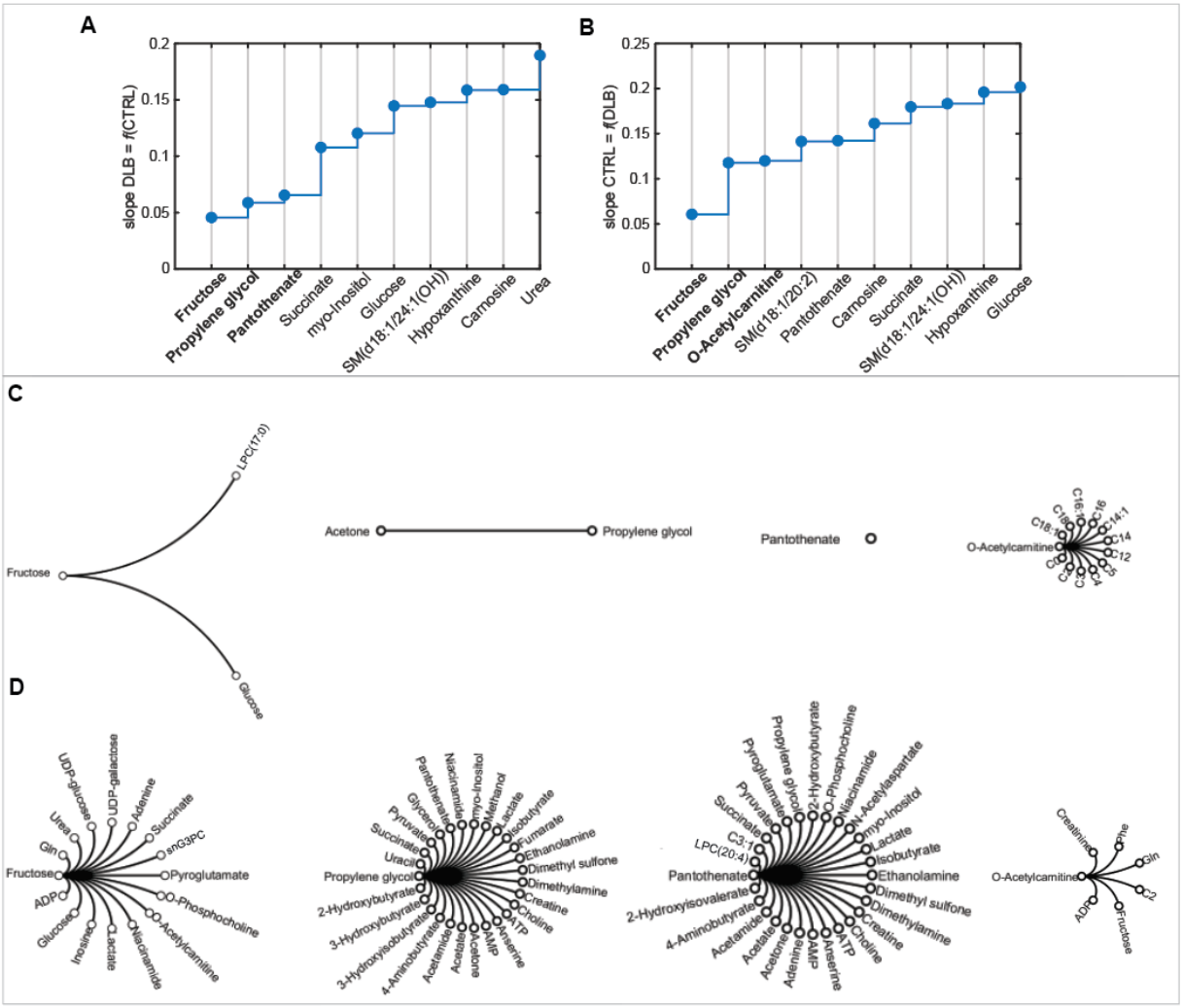


Figure 4. Comparison of Distance correlation analysis between control and DLB brain highlighting the metabolites with the largest interaction network change between the two groups. A. Metabolites with the largest change in the slope of the linear regression model between all correlations of each metabolite with

all the other metabolites calculated for CTRL and DLB groups; B. Same as A except using DLB as the variable and CTRL as the functionally dependent factor in the linear regression analysis; C. Networks for the four most altered metabolites with correlation threshold of 0.6 and $p<0.01$ in the CTRL group; D. Networks for the four most affected metabolites for the DLB cohort with correlation threshold of 0.6 and $p<0.01$. Order of metabolites in C and D is fructose, propylene glycol, pantothenate, *O*-acetylcarnitine.

Based on distance correlation network analysis (Figure 4), the fructose network changes from only a minimal number of significant correlations in the control group to a much larger number of significant edges in DLB suggesting a much more significant role for fructose in brain of DLB patients. Similarly, the second and the third most significantly altered metabolic networks (propylene-glycol and pantothenate metabolism) show more correlations with other metabolites in the DLB cohort. To the best of our knowledge, endogenous propylene-glycol has not been linked to dementia previously, possibly due to its limited concentration change. However, several pathways related to propylene-glycol have been reported as associated with dementia. On the other hand, pantothenate, i.e. vitamin B5, has been previously proposed as significantly changed in different types of dementias[27,26] . Based on distance correlation network analysis, pantothenate goes from no significant correlations in the control group to a number of significant correlations in the DLB patient group. Strong correlation partners of pantothenate include a number of metabolites that were shown to have significant concentration changes (Figure 1) as well as changes in clustering (Figure 2) and network correlation partners (Figure 4). *O*-acetylcarnitine has a network shift from significant correlations only with carnitines in the control cohort to correlations with ADP, fructose, and creatinine indicating a new associationin energy metabolism.

3.4 Partial correlation analysis

Partial correlation analysis indicates the effect of the removal of correlations for one of the partners on the correlations with other features there-by providing further information about likelihood of direct vs. indirect correlations. Partial correlation, using Spearman's correlation as the base model was performed for fructose, propylene-glycol, pantothenate and *O*-acetylcarnitine.

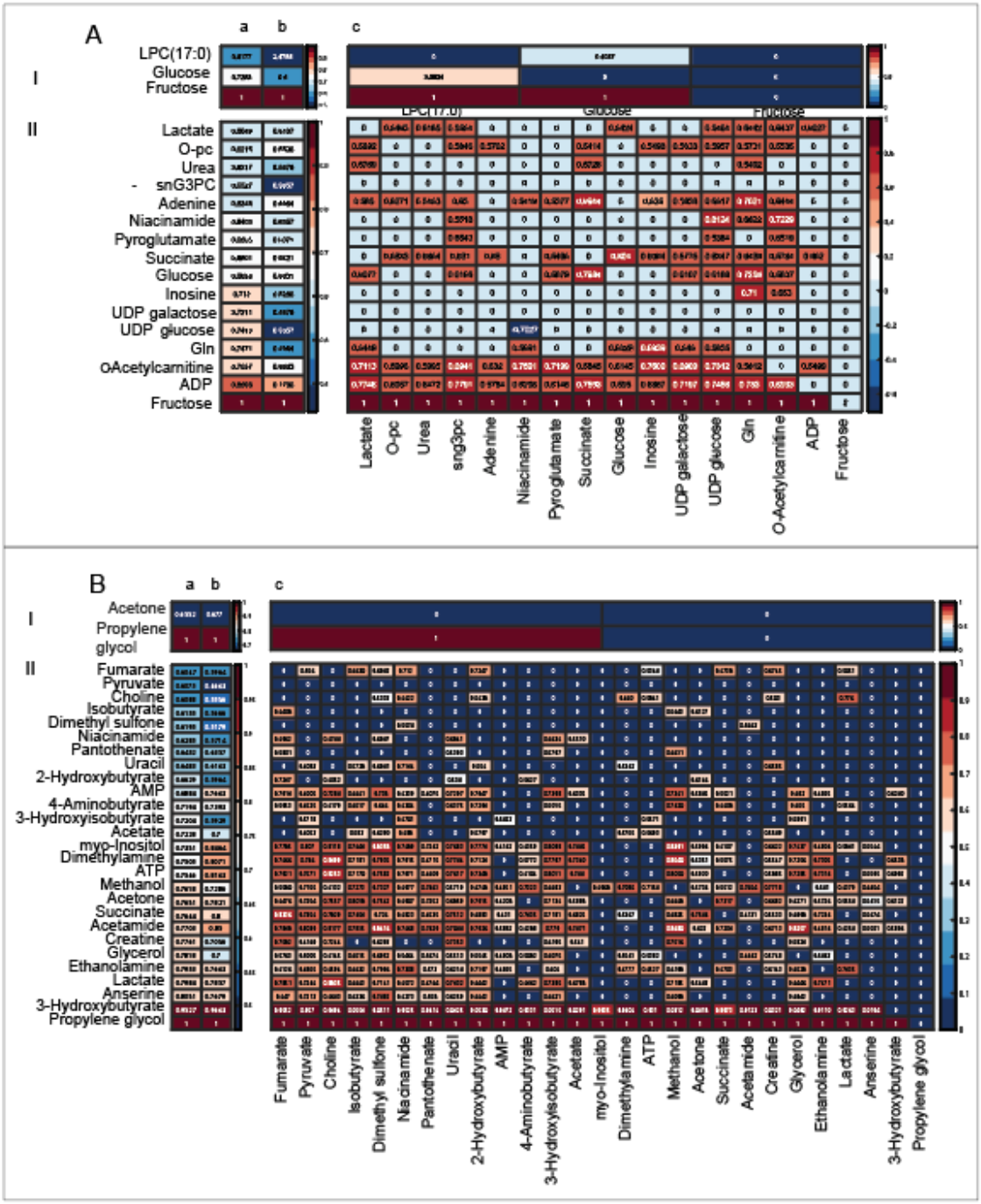


Figure 5. Distance correlation (a) Spearman (b) and Partial correlation (c) analysis of the most significant distance correlation partners of fructose for control (I) and DLB cohort (II) for major correlation partners of fructose (A) and propylene glycol (B). snG3PC – sn-glycero-3-phosphocholine.

In the control group (Figure 5I) fructose only has significant correlations with glucose and LPC(17:0) and, based on partial correlation analysis, the interactions are independent of each other. Conversely, if we remove the strongest correlator with fructose in DLB, ADP, we see removal of significant correlations with

all the other metabolites with exception of *O*-acetylcarnitine, succinate and lactate. This suggests that the correlation of fructose with the majority of metabolites in the DLB group results from the effect of fructose on ADP metabolism in this group.

The presentation of propylene-glycol as one of the metabolites with the most significant correlation change between control and DLB cohort is unexpected considering its very minor role in metabolism.

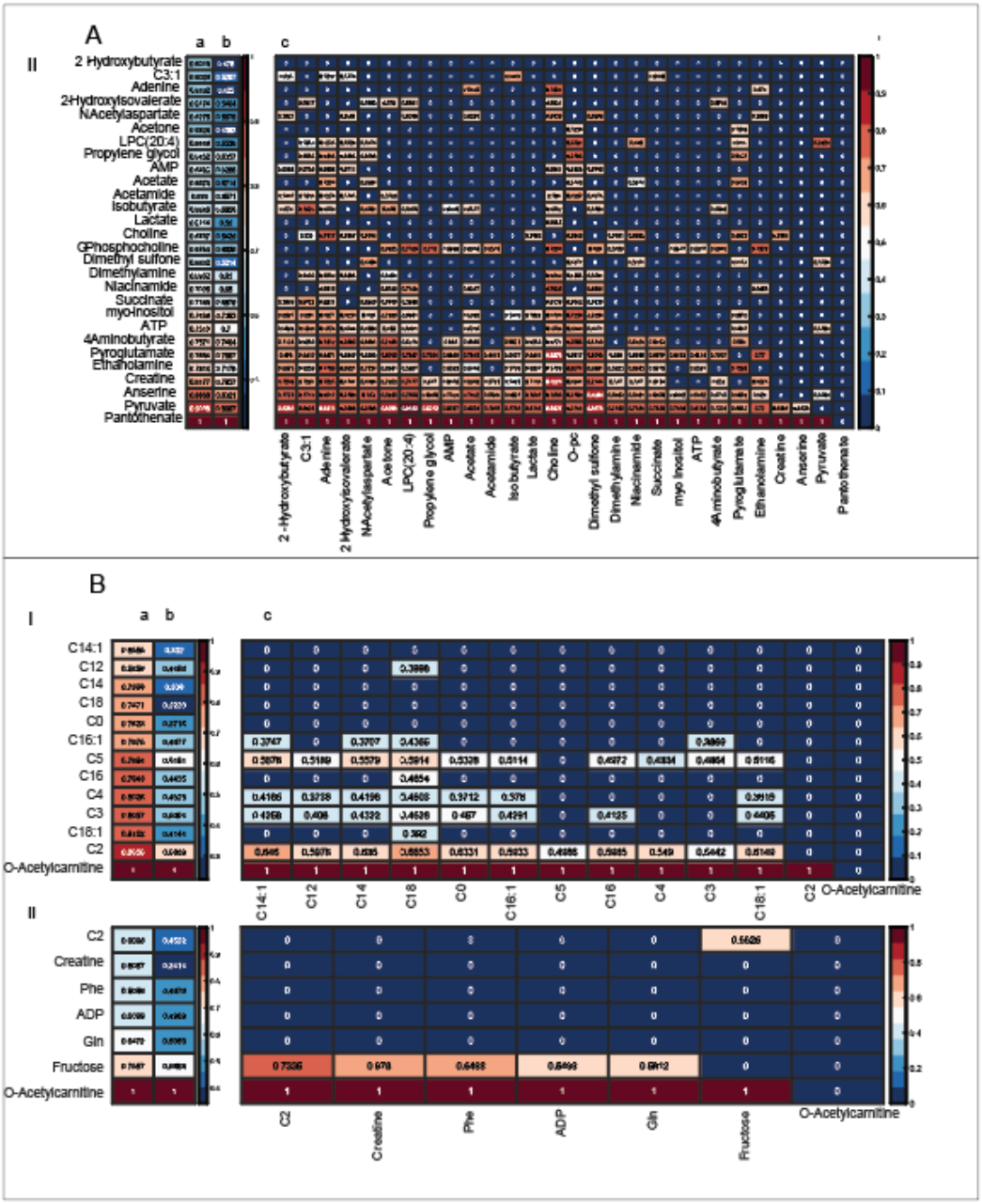


Figure 6. Distance correlation (a), Spearman (b) and Partial correlation (c) analysis of the most significant distance correlation partners of propylene-glycol for control (I) and DLB cohort (II) for major correlation partners of pantothenate (A) and *O*-acetylcarnitine (B).

In both control and DLB groups, propylene glycol has a strong correlation with acetone. In the DLB cohort propylene glycol has number of additional strong correlations with several metabolites that are involved in propanoate metabolism including 3-hydroxybutyrate, succinate, acetate and methanol. Within the propanoate metabolic pathway propylene glycol is metabolised into lactate, acetate, and pyruvate and significant correlations with these metabolites are reported.

The other correlation partners of propylene-glycol include pantothenate (vitamin B5), anserine, and dimethyl sulfone. However, the most significant partner of propylene-glycol is myo-inositol, with exclusions of its correlations resulting in the removal of almost all correlations for propylene-glycol as well. To a lesser extent, this is also achieved through correlations with dimethylamine and ATP.

Pantothenate, has no significant correlations in the control group and partial correlations are only shown for the DLB cohort (Figure 6). Major partners include creatine, anserine and pyruvate, once again with their correlation exclusion leading to the removal of the majority of correlations between pantothenate and its metabolic partners.

Finally, *O*-acetylcarnitine has a major change in its correlation network between control and DLB, where in control samples it is strongly correlated only with other acylcarnitines while in DLB samples it has strong correlations with fructose, glutamine, ADP, phenylalanine and creatine. Strong correlations between *O*-acetylcarnitine partners is demonstrated by the removal of correlations in the network in DLB through removal of partial correlations for each partner (Figure 6).

4. Discussion

Dementia with Lewy Bodies is a progressive and terminal disease for which there are currently no disease-modifying therapies. Furthermore, this disease can only be diagnosed based on its clinical presentation, as no biomarkers exists that will allow for early detection and commencement of therapy, and as such is often misdiagnosed. Earlier detection and proper stratification would allow for the matching of patients with specific and directed therapies based on an identified and mechanistically-understood metabolic dysregulation. Small-molecule metabolites and lipids are at the intersection of a cell or tissue's genetic background and environmental history. In particular, the levels of a specific metabolite or lipid within a cell or tissue are a consequence of its physiological, developmental, and pathological state[34]. These levels can be used as a reflection of specific phenotypes[35]. Due to the sheer number of species and range of concentrations of lipids, only looking at the concentration of each lipid may be short-sighted. Analysis of the metabolic network differences between patient groups is needed in order to determine functional changes between metabolites and across metabolic pathways and networks. Feature selection through statistical or machine learning approaches focuses on the determination of features that are most relevant for sample classification, thus not providing information about the functional changes leading to observed concentration differences. The development of interaction networks is a major step in understanding the role of metabolites in any given disease pathophysiology and the relevance of metabolic markers for diagnosis. Data-driven metabolic networks can be obtained through correlation or classification methods. Small sample size, incomplete knowledge of all the steps in a metabolic pathway, as well as sparse coverage of metabolites presents a unique challenge for network derivation. Here, we present metabolic networks for metabolites and lipids in the brain of DLB patients and matching controls as determined using signed distance correlation analysis. Metabolic network is explored using FCM clustering and partial correlation analysis. Our analysis has revealed major changes in metabolism of fructose, propylene-glycol, pantothenate and *O*-acetylcarnitine, concentration change for *O*-phosphocholine, *sn*-glycero-3-phosphocholine, putrescine, PC(O-34:0) and hydroxyvalerylcarnitine (C5-OH (C3-DC-M)) as well as a major overall change in network topology for a number of metabolic families in DLB.

Notably, fructose's correlation network goes from significant correlations only with glucose and LPC(17:0) in the control group to many significant edges in DLB. The role of fructose and its changing

metabolism has been previously hypothesized for different types of dementias and is considered as a risk factor in dementias[36,42]. The link between either endogenous or dietary fructose or ADP through the purine degradation pathway was suggested for Alzheimer's disease[36] and we report its strong correlation to DLB using our distance and partial correlation analysis (Figure 4-6). The purine degradation pathway is induced by fructose through its rapid depletion of cellular ATP levels and activation of AMP deaminase (AMPD), eventually leading to the production of uric acid. Changes in the level of uric acid have been previously linked to the risk of dementia, with low levels associated with Parkinson's disease[36,31] and high levels of serum uric acid linked to vascular or mixed dementia and to a lesser extent AD[41].

Distance correlation analysis shows major differences in correlation partners for fructose in patients with DLB compared to healthy controls, including significant correlations with several other metabolites that are part of purine metabolism including ADP, adenine, inosine and urea. Significant changes, especially in certain acylcarnitine and LPC levels, have been previously associated with a high fructose diet[5]. A number of metabolites in the fructose network are also part of lactose synthesis and galactose metabolism. These include glucose, UDP glucose, UDP galactose, ADP and fructose. Fructose production from glucose through the polyol pathway has been previously observed in different tissues including brain[28].The production of fructose through this pathway stimulates triglyceride and uric acid accumulation and is hypothesized to be a relevant factor in the development of metabolic syndrome. Up-regulation of aldose reductase, a rate-limiting enzyme in this pathway, has been shown in aging and as a response to a number of dementia risk factors[36]. It is important to point out that the concentration of fructose does not show a significant difference between control and DLB cohorts, suggesting a significant change in the fructose metabolic pathway rather than its accumulation or depletion in these patients

It is important to emphasize that in the DLB group, removal of correlations with the strongest connection to fructose (ADP) leads to the elimination of correlations with all other metabolites except *O*-acetylcarnitine, succinate and lactate. This result suggests that correlation with most metabolites results from the effect of fructose, either endogenous from the polyol pathway or from the diet, on ADP metabolism. This indicates that in DLB patients, as was previously hypothesized for AD[36], fructose induces the purine degradation pathway by increasing ATP consumption, resulting in the production of uric acid. Adenine, inosine, urea, and glutamine all partake in purine metabolism. However, *O*-acetylcarnitine and succinate are part of the Oxidation of Branched Chain Fatty Acids pathway. Profound effects of fructose on cell metabolism and in particular mitochondrial function have been reported previously[43] and include divergent effects of fructose and glucose on mitochondrial function and fatty acid oxidation.

In both control and DLB cohorts, propylene glycol has a strong correlation with acetone which may indicate a link via propanoate metabolism that has been previously linked to aging and AD[38]. Propylene glycol is metabolised to lactate, acetate, and pyruvate and significant correlations with these metabolites are presented in the DLB group. In fact, propylene glycol as well as several of its correlation partners, i.e., succinate, acetate, lactate, acetone, pyruvate, 2-hydroxybutyrate are all involved in propanoate metabolism (based on KEGG map00640 pathway list). Additionally, through β-Alanine metabolism, propylene glycol is once again linked to the propanoate metabolic pathways which includes anserine (part of β-Alanine metabolism) and pantothenate, both of which correlated strongly with propylene-glycol in DLB. Several metabolites from the propanoate pathway have previously been reported as significantly different in saliva of dementia patients[32] and this pathway has been shown to be affected in AD patients[39]. Further, propylene-glycol is used as a solvent for several intravenously administered drugs including lorazepam and diazepam, both of which are used to treat symptoms associated with DLB. Its correlation needs further examination as we are unsure if the patients studied herein were treated with said drugs. Moreover, pantothenate (vitamin B5), anserine, dimethyl sulfone, could also be considered results of therapy either as drugs or supplements[30]. For example, anserine is a supplement used to improve symptoms of dementia; dimethyl sulfone is an anti-pain, inflammatory and osteoarthritis drug (https://go.drugbank.com/drugs/DB14090) and pantothenate or vitamin B5 is a recommended supplement for dementia patients.

Pantothenate shows the third most significant network change with strong correlations to propylene glycol, anserine and dimethyl sulfone. Pantothenate has a significant metabolic role with involvement in β-alanine metabolism and CoA biosynthesis. In the control group, pantothenate has no significant correlation partners at correlation level over 0.6 and p-value below 0.01. In the DLB group, pantothenate gains several strong correlation partners, including 4-aminobutanoate, a by-product of one of the steps of β-alanine synthesis, as well as anserine, suggesting its role in metabolism changes in DLB patients.

There are several strengths of the present study. These include the standardized brain collections, the use of age- and sex-matched controls of this unique cohort, a combined quantitative metabolomic approach delivering the most comprehensive metabolite coverage of the DLB brain metabolome and implementation of novel, analytical approaches within the realm of DLB metabolomics. The small sample size of this unique cohort is a major limitation; however, one should be cognisant of the problems in obtaining such a specialized and well characterized post-mortem brain cohort. Through the examination of well characterized samples, even at this sample size, we can develop models of high diagnostic accuracy. It would be worthwhile to expand the analysis carried out here to other brain regions to assess the wider metabolic disturbances in the DLB brain. Brain tissue is far from an ideal matrix for discovering dementia biomarkers, and it will be necessary to validate our findings in more accessible, non-invasive biomatrices such as blood serum/plasma as evidenced by Varma et al. for AD[44]. Another major limitation which may have provided some additional, useful insight to our exploratory study is the lack of a detailed medical report with important information regarding medications and supplements.

In this study, we introduce a novel approach of signed distance in addition to the novel application of partial correlation analysis and FCM clustering for the study of DLB. The application of these novel analytical approaches on such a unique metabolomics data set highlights fructose, propylene-glycol and pantothenate acid as well as their associated metabolic pathways as key factors associated with DLB pathology. Our findings have the potential to provide new insight into the pathophysiology of DLB which could lead to the development of novel therapeutics and novel diagnostic methods capable of accurately discriminating DLB from control brain with a high degree of accuracy. Although brain tissue is not an ideal matrix for developing biomarker tests, understanding of the major changes in brain metabolism/biochemistry are crucial for future treatment development. Additionally, methods presented in this work are made available for future use for the analysis of functional metabolic markers of DLB in more accessible tissues.

Research in Context

1. Systematic review: We performed extensive literature review using publication databases (e.g. PubMed) and focusing on dementias, DLB and role of metabolism and metabolites in the disease development and progression.
2. Interpretation: Using a novel approach for analysis of metabolic network we have shown major changes in metabolism of fructose, o-acetylcarnitine, propylene glycol and pantothenate in brain tissue of DBL patients
3. Future directions: Focus on metabolic network changes rather than only metabolite concentration alterations will bring novel insight into the DLB development and provide new therapeutic targets.

Data Availability

All data is available at: https://doi.org/10.4224/40002692

Institutional Review Board Statement

This study was approved by the Beaumont Health System's Human Investigation Committee (HIC No.: 2018-387).

Informed Consent Statement

Not applicable.

Competing Interest Statement

S.F.G. is supported by grants from the National Institute of Neurological Disorders and Stroke (1R01NS110838- 01A1), the National Institutes on Aging (1R21AG067083-01) and the Michael J. Fox Foundation (MJFF16201). All other authors declare no perceived or true conflicts of interest.

Author Contributions

Conceptualization, M.C.C., S.F.G., S.A.L.B.; Methodology, M.C.C., A.Y., S.A., S.A.L.B. and S.F.G.; Formal Analysis, M.C.C.; Experimental Analysis: A.Y.; S.A.; S.V.; S.F.G.; Clinical Samples: B.M.; P.P.; P.G.K.; M.E.M.; B.D.G.; Software development, D.S., A.S. and M.C.C.; Supervision, S.F.G., M.C.C., S.A.L.B.; Writing—Original Draft Preparation, M.C.C; Writing—Review and Editing, M.C.C., S.F.G. I.A., T.N-T.; B.D.G.; A.Y.; All authors have read and agreed to the published version of the manuscript.

Acknowledgements

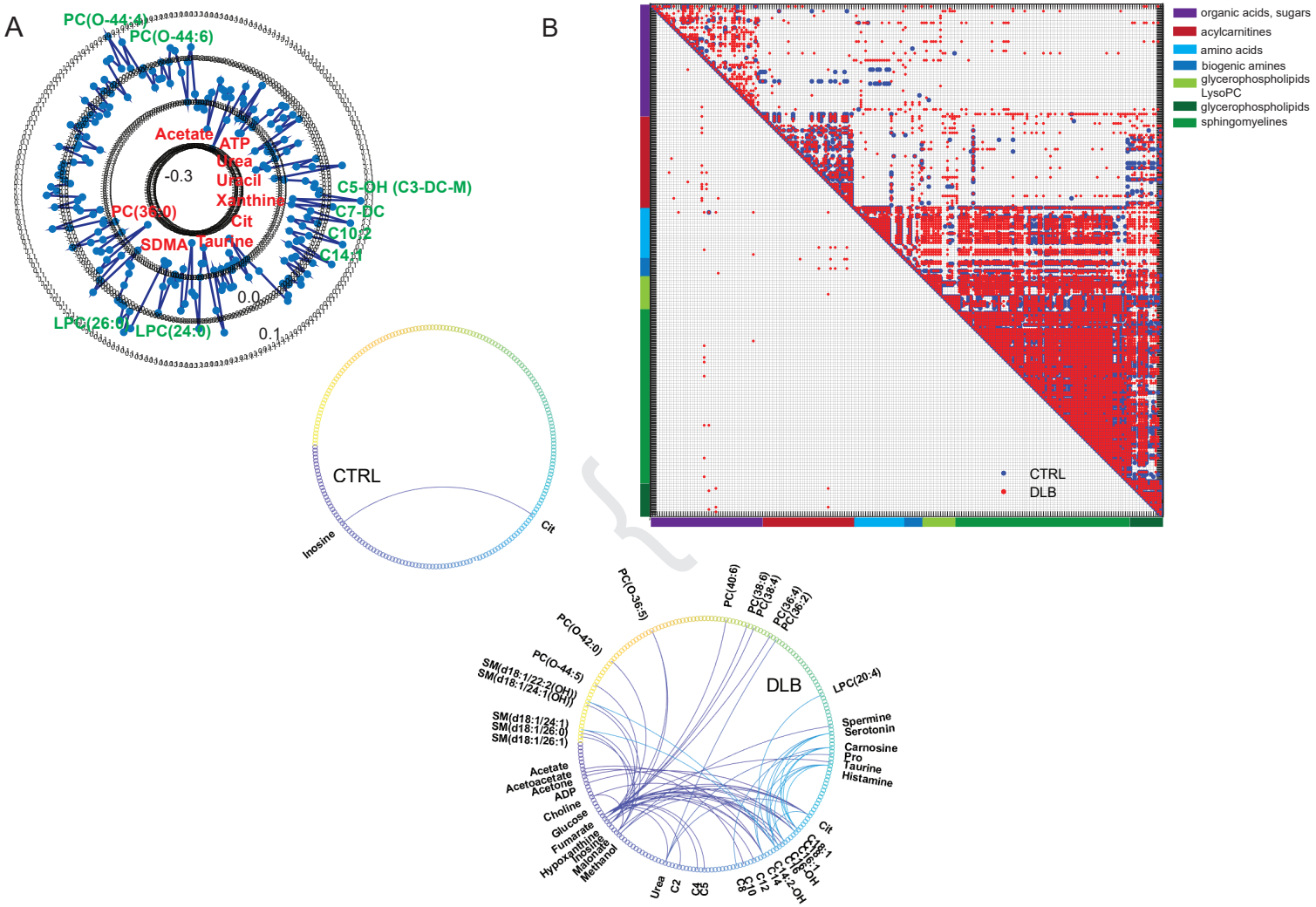
We thank the Brains for Dementia Research (BDR) initiative, a brain bank network funded by Alzheimer's Brain Bank UK (ABBUK), a charity co-funded by Alzheimer's Research UK (ARUK), and the Alzheimer's Society, for graciously providing the tissue samples for this study. This work was partly funded by the generous contribution made by the John and Marilyn Bishop Charitable Foundation and the Fred A. & Barbara M. Erb Foundation. Most importantly, we would like to thank all patients and volunteers for their invaluable and selfless contribution to this research.

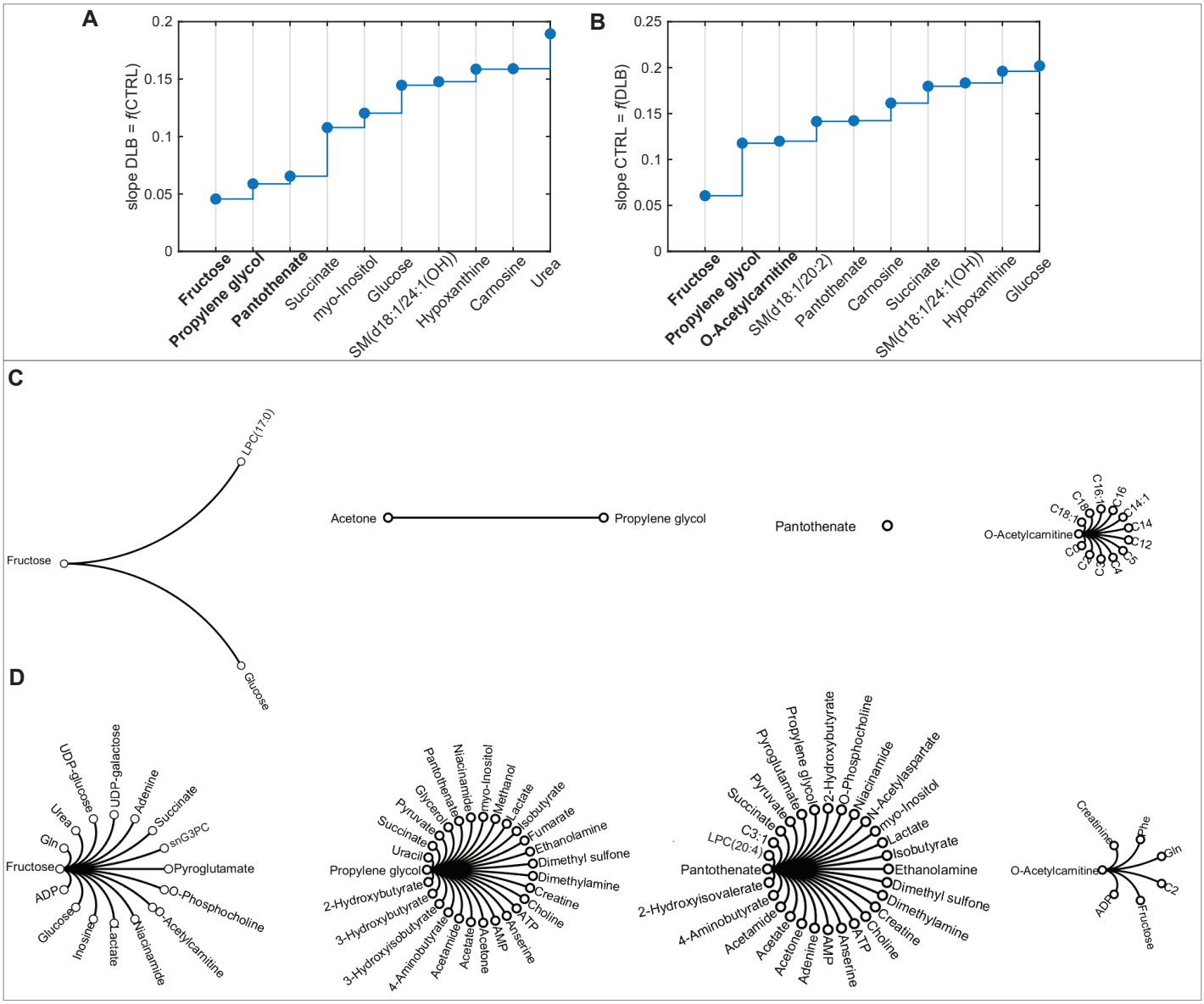
References

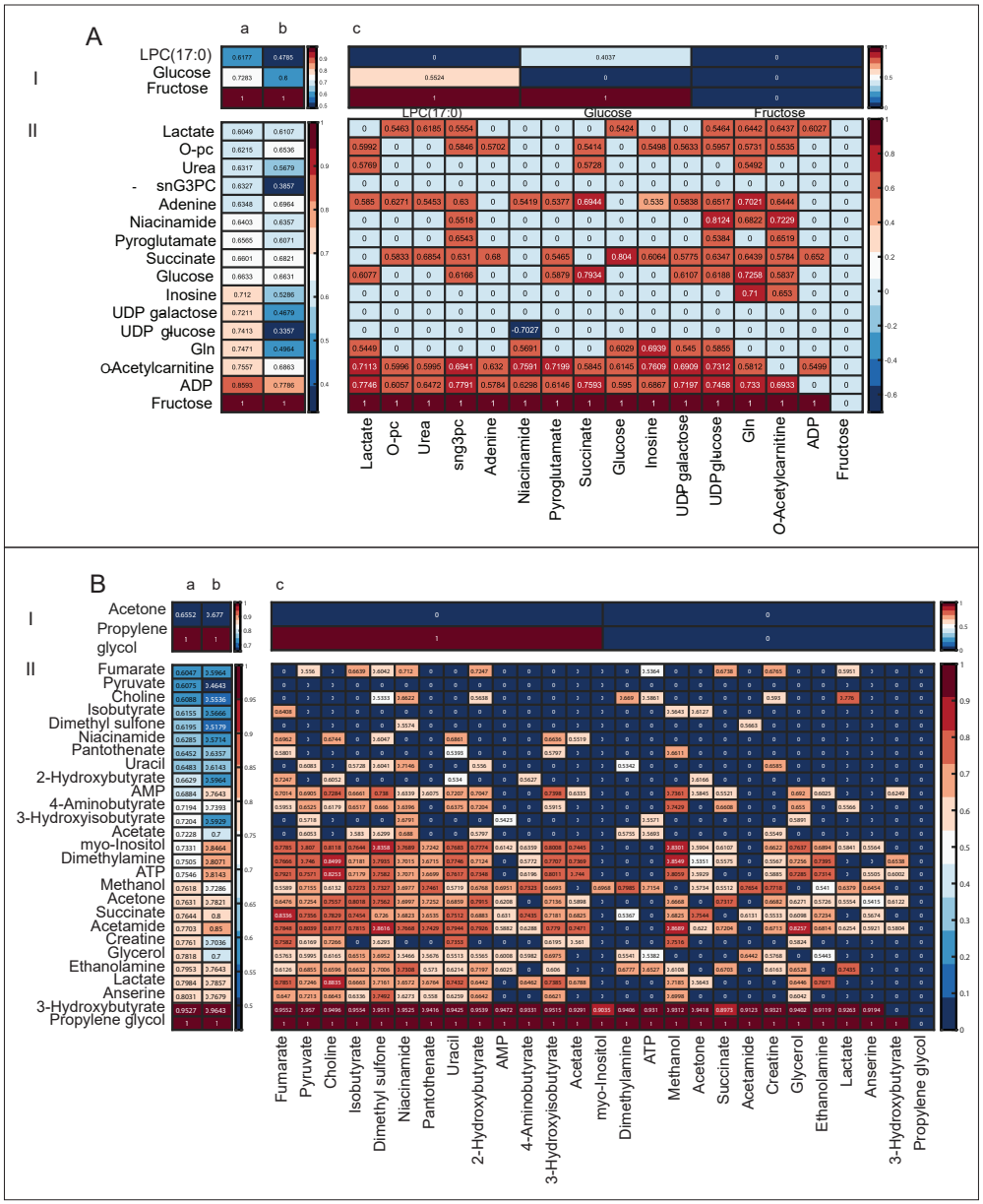
1. Arnaoutoglou NA, O'Brien JT, Underwood BR. Dementia with Lewy bodies - from scientific knowledge to clinical insights. *Nat Rev Neurol.* 2019; **15**:103-112.
2. McKeith, I.G. et al. Research criteria for the diagnosis of prodromal dementia with Lewy bodies *Neurology* 2020; **94**: 743-755.
3. Outeiro TF, Koss DJ, Erskine D, Walker L, Kurzawa-Akanbi M, Burn D, Donaghy P, et al.. Dementia with Lewy Bodies: An Update and Outlook. *Molecular Neurodegeneration*, 2019; 14: 5.
4. Taylor JP, McKeith IG, Burn DJ, Boeve BF, Weintraub D, Bamford C, Allan LM, Thomas AJ, O'Brien JT. New evidence on the management of Lewy body dementia. *Lancet Neurol.* 2020 **19**:157-169.
5. Garcia-Esparcia, Paula, Irene López-González, Oriol Grau-Rivera, María Francisca García-Garrido, Anusha Konetti, Franc Llorens, Saima Zafar, et al.. "Dementia with Lewy Bodies: Molecular Pathology in the Frontal Cortex in Typical and Rapidly Progressive Forms." *Frontiers in Neurology* 2017; **8**: 89.
6. Leskelä S, Hoffmann D, Rostalski H, Huber N, Wittrahm R, Hartikainen P, Korhonen V, et al. "FTLD Patient–Derived Fibroblasts Show Defective Mitochondrial Function and Accumulation of p62. *Molecular Neurobiology*.2021, **58**: 5438-5458.
7. Akyol S, Yilmaz A, Oh KJ, Ugur Z, Aydas B, McGuinness B, Passmore P, et al.. Evidence That the Kennedy and Polyamine Pathways Are Dysregulated in Human Brain in Cases of Dementia with Lewy Bodies." *Brain Research* 2020; **1743**: 146897.
8. Fanning S, Selkoe D, Dettmer U. Parkinson's disease: proteinopathy or lipidopathy? *NPJ Parkinsons Dis.* 2020; **3**: 3.
9. Alecu, I. & Bennett, SAL. Dysregulated lipid metabolism and its role in α-synucleinopathy in Parkinson's disease. *Front. Neurosci*. 2019; **13**, 328.
10. Plotegher N, Bubacco L, Greggio E, Civiero L. Ceramides in Parkinson's Disease: From Recent Evidence to New Hypotheses. *Front Neurosci.* 2019;**13**:330
11. Jeong H, Tombor B, Albert R, Oltvai ZN, Barabási AL. The large-scale organization of metabolic networks. *Nature*. 2000; **407**: 651-4.
12. Ravasz E, A. Somera D, Mongru Z, Oltvai AL, Barabàsi Hierarchical organization of modularity in metabolic networks. *Science*, 2002; **297**: 1551-1555
13. Robinson L, Kocabaş P, Wang H, Cholley P.-E., Cook D, Nilsson A, Anton M, Ferreira R, Domenzain I, Billa V, et al. An atlas of human metabolism. Sci. Signal., 13 (2020), p. eaaz1482,
14. Beguerisse-Díaz, M., Bosque, G., Oyarzún, D. et al. Flux-dependent graphs for metabolic networks. *NPJ Syst Biol Appl* 2018; **4**: 32.
15. Székely, Gábor J., Maria L. Rizzo, and Nail K. Bakirov. 2007. "Measuring and Testing Dependence by Correlation of Distances. *The Annals of Statistics*, 2007, **35**: 2769-2794..
16. Pardo-Diaz, J, L. V. Bozhilova, M. Beguerisse-Díaz, P. S. Poole, C. M. Deane, and G. Reinert.. "Robust Gene Coexpression Networks Using Signed Distance Correlation." *Bioinformatics 2021; **37**: 1982-1989.*.
17. Cuperlović-Culf M, Belacel N, Culf AS, Chute IC, Ouellette RJ, Burton IW, Karakach TK, Walter JA, NMR Metabolic Analysis of Samples Using Fuzzy K-Means Clustering. *Magnetic Resonance in Chemistry: MRC* 2009; **47** Suppl 1: S96–104.
18. Belacel N, Cuperlović-Culf M, Laflamme M, Ouellette R. "Fuzzy J-Means and VNS Methods for Clustering Genes from Microarray Data." *Bioinformatics* 2004; **20**: 1690–1701.
19. Graham SF, Pan X, Yilmaz A, Macias S, Robinson A, Mann D, Green BD. Targeted biochemical profiling of brain from Huntington's disease patients reveals novel metabolic pathways of interest. *Biochim Biophys Acta Mol Basis Dis*. 2018; **1864**: 2430-2437.
20. Graham SF, Kumar PK, Bjorndahl T, Han B, Yilmaz A, Sherman E, Bahado-Singh RO, Wishart, D. Mann, B. Green Metabolic signatures of Huntington's disease (HD): (1)H NMR analysis of the polar metabolome in post-mortem human brain. *Biochim. Biophys. Acta,* 2016; **1862**: 1675-1684
21. Graham SF, Kumar P, R.O. Bahado-Singh, A. Robinson, D. Mann, BD. Green Novel metabolite biomarkers of Huntington's disease (HD) as detected by high resolution mass spectrometry *J. Proteome Res.,* 2016; **15**: 1592-1601
22. Dunn, JC. "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters." *Journal of Cybernetics*.1973; **3**: 32-57.
23. Lionelle, A. *Unsupervised ReliefF*.2005; Colorado State Univ.

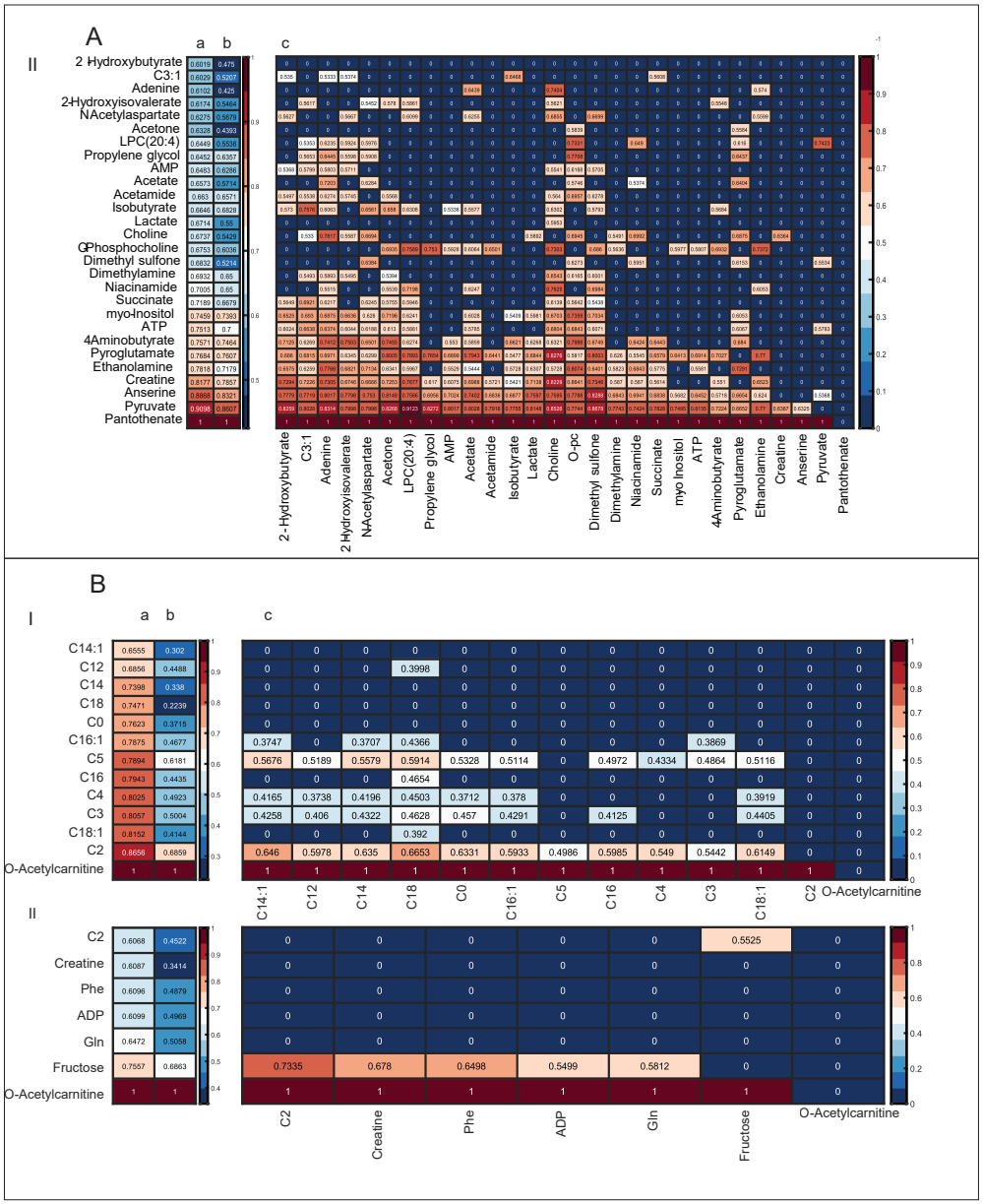
24. Tibshirani, R. “Regression Shrinkage and Selection via the Lasso.” *Journal of the Royal Statistical Society.* Series B, 1996, **58**: 267–288.
25. Pang Z, Chong J, Zhou G, Anderson de Lima Morais D, Le Chang, Barrette M, Gauthier C, Jacques P, Li S, Xia J.. “MetaboAnalyst 5.0: Narrowing the Gap between Raw Spectra and Functional Insights.” *Nucleic Acids Research* 2021; **49** (W1): W388–96.
26. Patassini S, Begley P, Xu J, Church S, Kureishy N, Reid S, Waldvogel H, et al.. “Cerebral Vitamin B5 (D-Pantothenic Acid) Deficiency as a Potential Cause of Metabolic Perturbation and Neurodegeneration in Huntington’s Disease.” *Metabolites*, 2019; **9**: 113.
27. Xu J, Begley P, Church SJ, Patassini S, McHarg S, Kureishy N, Hollywood KA, et al. “Elevation of Brain Glucose and Polyol-Pathway Intermediates with Accompanying Brain-Copper Deficiency in Patients with Alzheimer’s Disease: Metabolic Basis for Dementia.” *Scientific Reports* 2016; **6**: 27524.
28. Andres-Hernando, Ana, Richard J. Johnson, and Miguel A. Lanaspa. Endogenous Fructose Production: What Do We Know and How Relevant Is It? *Current Opinion in Clinical Nutrition and Metabolic Care 2019;* **22**: 289.
29. Bezdek JC. “Modified Objective Function Algorithms.” *Pattern Recognition with Fuzzy Objective Function Algorithms*.1981; **4:** 508-516.
30. Ding Q, Tanigawa K, Kaneko J, Totsuka M, Katakura Y, Imabayashi E, Matsuda H, Hisatsune T. Anserine/Carnosine Supplementation Preserves Blood Flow in the Prefrontal Brain of Elderly People Carrying APOE e4. *Aging and Disease* 2018; **9**: 334.
31. Ellmore TM, Suescun J, Castriotta RJ, Schiess MC. “A Study of the Relationship Between Uric Acid and Substantia Nigra Brain Connectivity in Patients With REM Sleep Behavior Disorder and Parkinson’s Disease.” *Frontiers in Neurology* 2020; **815**.
32. Figueira J, Jonsson P, Adolfsson AN, Adolfsson R, Nyberg L, Öhman A. NMR Analysis of the Human Saliva Metabolome Distinguishes Dementia Patients from Matched Controls. *Molecular bioSystems* 2016; **12**: 2562–71.
33. Gonzalez-Granda A, Damms-Machado A, Basrai M, Bischoff SC. 2018. “Changes in Plasma Acylcarnitine and Lysophosphatidylcholine Levels Following a High-Fructose Diet: A Targeted Metabolomics Study in Healthy Women.” *Nutrients,* 2018*;* **10**: 1254.
34. Fiehn O. Metabolomics—The link between genotypes and phenotypes. *Plant Molecular Biology*, 2002; **48**:155-171.
35. Shamim A, Mahmood T, Ahsan F, Kumar A, Bagga P. Lipids: An insight into the neurodegenerative disorders*. Clinical Nutrition Experimental*, 2018; **20**:1-19.
36. Johnson RJ, Gomez-Pinilla F, Nagel M, Nakagawa T, Rodriguez-Iturbe B, Sanchez-Lozada LG, Tolan DR, Lanaspa MA.. “Cerebral Fructose Metabolism as a Potential Mechanism Driving Alzheimer’s Disease.” *Frontiers in Aging Neuroscience* 2020; **560865**.
37. Kaneko J, Enya A, Enomoto K, Ding Q, Hisatsune T.. “Anserine (beta-Alanyl-3-Methyl-L-Histidine) Improves Neurovascular-Unit Dysfunction and Spatial Memory in Aged AβPPswe/PSEN1dE9 Alzheimer’s-Model Mice.” *Scientific Reports* 2017, **7**: 1–12.
38. Killingsworth J, Sawmiller D, Shytle RD. “Propionate and Alzheimer’s Disease.” *Frontiers in Aging Neuroscience* 2020; **12**. 580001.
39. Kong, Y., Wu J, Yuan L. “MicroRNA Expression Analysis of Adult-Onset Drosophila Alzheimer’s Disease Model.” *Current Alzheimer Research*, 2014; **11**: 815.
40. Kucera M, Isserlin R, Arkhangorodsky A, Bader GB. AutoAnnotate: A Cytoscape App for Summarizing Networks with Semantic Annotations. *F1000Research* 2016; **5**: 20161717.
41. Latourte A, Soumaré A, Bardin T, Perez-Ruiz F, Debette S, Richette P. Uric Acid and Incident Dementia over 12 Years of Follow-up: A Population-Based Cohort Study. *Annals of the Rheumatic Diseases* 2018; **77**: 328–35.
42. Xu J, Patassini S, Begley P, Church S, Waldvogel HJ, Faull RLM, Unwin RD, Cooper GJS. Cerebral Deficiency of Vitamin B5 (d-Pantothenic Acid; Pantothenate) as a Potentially-Reversible Cause of Neurodegeneration and Dementia in Sporadic Alzheimer’s Disease. *Biochemical and Biophysical Research Communications* 2020; **527**: 676–81.
43. Softic S, Meyer JG, Wang G, Gupta MK et al. Dietary Sugars Alter Hepatic Fatty Acid Oxidation via Transcriptional and Post-translational Modifications of Mitochondrial Proteins, *Cell Metabolism,* 2019, **30**: 735-753.e4.
44. Varma VR, Oommen AM, Varma S, et al. Brain and blood metabolite signatures of pathology and

progression in Alzheimer disease: A targeted metabolomics study. *PLoS Med.* 2018; **15**: e1002482.


A
PC(O-44:4)
PC(O-44:6)
Acetate
ATP
Urea
Uracil
Xanthine
Cit
Taurine
SDMA
PC(36:0)
C5-OH (C3-DC-M)
C7-DC
C10:2
C14:1
LPC(26:0)
LPC(24:0)
-0.3
0.0
0.1
B
organic acids, sugars
acylcarnitines
amino acids
biogenic amines
glycerophospholipids
LysoPC
glycerophospholipids
sphingomyelines
CTRL
DLB
CTRL
Inosine
Cit
DLB
PC(O-36:5)
PC(40:6)
PC(38:6)
PC(38:4)
PC(36:4)
PC(36:2)
PC(O-42:0)
PC(O-44:5)
SM(d18:1/22:2(OH))
SM(d18:1/24:1(OH))
SM(d18:1/24:1)
SM(d18:1/26:0)
SM(d18:1/26:1)
Acetate
Acetoacetate
Acetone
ADP
Choline
Glucose
Fumarate
Hypoxanthine
Inositol
Malonate
Methanol
Urea
C2
C4
C5
C8:0
C12
C14:2-OH
C16:1-OH
C18:1
Cit
Histamine
Taurine
Pro
Carnosine
Serotonin
Spermine
LPC(20:4)


A
slope DLB = f(CTRL)
0.2
0.15
0.1
0.05
0
Fructose
Propylene glycol
Pantothenate
Succinate
myo-Inositol
Glucose
SM(d18:1/24:1(OH))
Hypoxanthine
Carnosine
Urea
B
slope CTRL = f(DLB)
0.25
0.2
0.15
0.1
0.05
0
Fructose
Propylene glycol
O-Acetylcarnitine
SM(d18:1/20:2)
Pantothenate
Carnosine
Succinate
SM(d18:1/24:1(OH))
Hypoxanthine
Glucose
C
Fructose
LPC(17:0)
Glucose
Acetone
Propylene glycol
Pantothenate
O-Acetylcarnitine
C16:1
C16
C14:1
C18:1
C18
C14
C12
C10
C8
C6
C5
C4
C3
D
Fructose
UDP-glucose
UDP-galactose
Adenine
Succinate
Urea
Gln
snG3PC
Pyroglutamate
ADP
O-Phosphocholine
Glucose
Inosine
Lactate
Niacinamide
O-Acetylcarnitine
Propylene glycol
Pantothenate
Niacinamide
myo-Inositol
Glycerol
Methanol
Pyruvate
Lactate
Succinate
Isobutyrate
Uracil
Fumarate
Ethanolamine
Dimethyl sulfone
2-Hydroxybutyrate
Dimethylamine
3-Hydroxybutyrate
Creatine
3-Hydroxyisobutyrate
Choline
4-Aminobutyrate
ATP
Acetamide
Anserine
Acetate
AMP
Acetone
Pantothenate
Propylene glycol
Pyroglutamate
2-Hydroxybutyrate
Pyruvate
O-Phosphocholine
Succinate
Niacinamide
C3:1
N-Acetylaspartate
LPC(20:4)
myo-Inositol
Lactate
Isobutyrate
Ethanolamine
2-Hydroxyisovalerate
Dimethyl sulfone
4-Aminobutyrate
Dimethylamine
Acetamide
Creatine
Acetate
Choline
Acetone
ATP
Adenine
Anserine
AMP
O-Acetylcarnitine
Creatinine
Phe
Gln
C2
ADP
Fructose


A
a
b
c
I
LPC(17:0)
Glucose
Fructose
II
Lactate
O-pc
Urea
snG3PC
Adenine
Niacinamide
Pyroglutamate
Succinate
Glucose
Inosine
UDP galactose
UDP glucose
Gln
O-Acetylcarnitine
ADP
Fructose
B
Acetone
Propylene glycol
Fumarate
Pyruvate
Choline
Isobutyrate
Dimethyl sulfone
Niacinamide
Pantothenate
Uracil
2-Hydroxybutyrate
AMP
4-Aminobutyrate
3-Hydroxyisobutyrate
Acetate
myo-Inositol
Dimethylamine
ATP
Methanol
Acetone
Succinate
Acetamide
Creatine
Glycerol
Ethanolamine
Lactate
Anserine
3-Hydroxybutyrate
Propylene glycol


A
II
a
b
c
2 Hydroxybutyrate
C3:1
Adenine
2-Hydroxyisovalerate
NAcetylaspartate
Acetone
LPC(20:4)
Propylene glycol
AMP
Acetate
Acetamide
Isobutyrate
Lactate
Choline
GPhosphocholine
Dimethyl sulfone
Dimethylamine
Niacinamide
Succinate
myo-Inositol
ATP
4Aminobutyrate
Pyroglutamate
Ethanolamine
Creatine
Anserine
Pyruvate
Pantothenate
B
I
C14:1
C12
C14
C18
C0
C16:1
C5
C16
C4
C3
C18:1
C2
O-Acetylcarnitine
II
C2
Creatine
Phe
ADP
Gln
Fructose
O-Acetylcarnitine

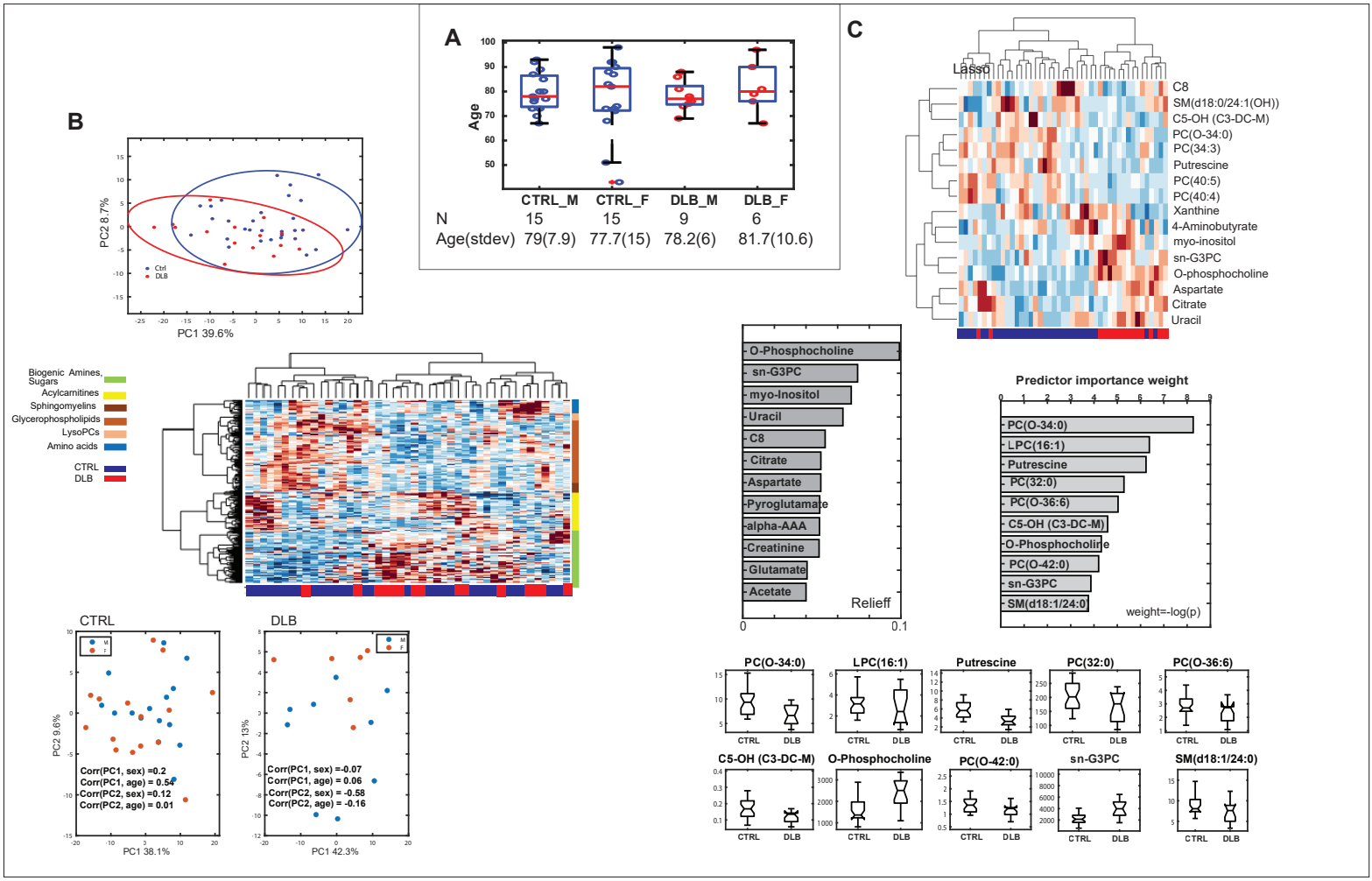
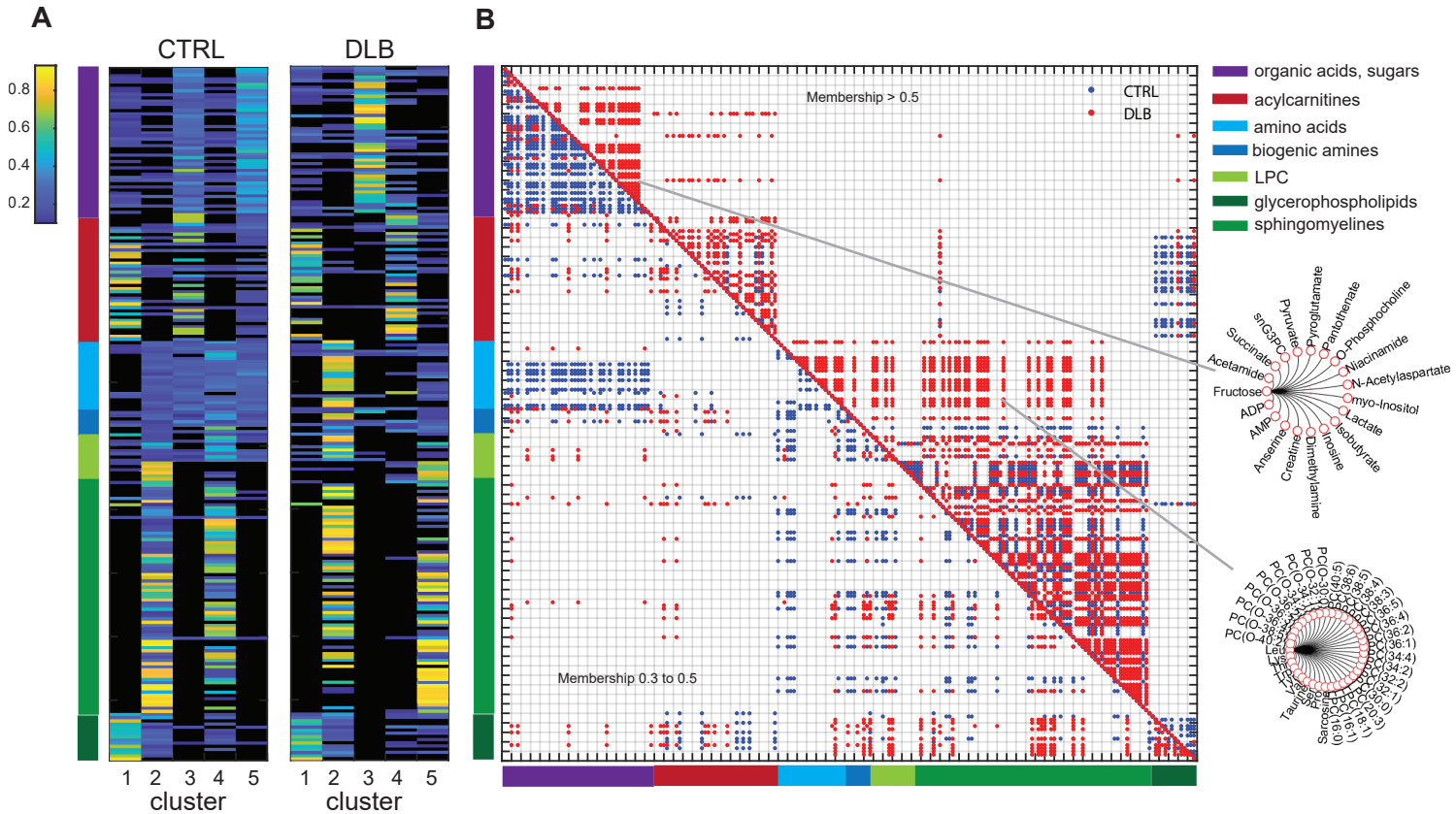

A
Age
CTRL_M CTRL_F DLB_M DLB_F
N 15 15 9 6
Age(stdev) 79(7.9) 77.7(15) 78.2(6) 81.7(10.6)
B
PC2 8.7%
PC1 39.6%
Ctrl
DLB
Biogenic Amines, Sugars
Acylcarnitines
Sphingomyelins
Glycerophospholipids
LysoPCs
Amino acids
CTRL
DLB
CTRL
DLB
PC2 9.6%
PC1 38.1%
Corr(PC1, sex) =0.2
Corr(PC1, age) = 0.54
Corr(PC2, sex) =0.12
Corr(PC2, age) = 0.01
PC2 13%
PC1 42.3%
Corr(PC1, sex) =-0.07
Corr(PC1, age) = 0.06
Corr(PC2, sex) = -0.58
Corr(PC2, age) = -0.16
C
Lasso
C8
SM(d18:0/24:1(OH))
C5-OH (C3-DC-M)
PC(O-34:0)
PC(34:3)
Putrescine
PC(40:5)
PC(40:4)
Xanthine
4-Aminobutyrate
myo-inositol
sn-G3PC
O-phosphocholine
Aspartate
Citrate
Uracil
O-Phosphocholine
sn-G3PC
myo-Inositol
Uracil
C8
Citrate
Aspartate
Pyroglutamate
alpha-AAA
Creatinine
Glutamate
Acetate
Relieff
0
0.1
Predictor importance weight
PC(O-34:0)
LPC(16:1)
Putrescine
PC(32:0)
PC(O-36:6)
C5-OH (C3-DC-M)
PC(O-42:0)
SM(d18:1/24:0)
weight=-log(p)
PC(O-34:0)
LPC(16:1)
Putrescine
PC(32:0)
PC(O-36:6)
C5-OH (C3-DC-M)
O-Phosphocholine
PC(O-42:0)
sn-G3PC
SM(d18:1/24:0)
CTRL
DLB

Figure 1


A
CTRL
DLB
0.8
0.6
0.4
0.2
1 2 3 4 5
cluster
1 2 3 4 5
cluster
B
Membership > 0.5
Membership 0.3 to 0.5
CTRL
DLB
organic acids, sugars
acylcarnitines
amino acids
biogenic amines
LPC
glycerophospholipids
sphingomyelines
Pyroglutamate
Pantothenate
O-Phosphocholine
Pyruvate
sn G3PC
Succinate
Niacinamide
Acetamide
N-Acetylaspartate
Fructose
myo-Inositol
ADP
Lactate
AMP
Isobutyrate
Anserine
Creatine
Dimethylamine
Inosine