Circulatory dietary and gut-derived metabolites predict preclinical Alzheimer's disease.

Emily Connell[1], Saber Sami[1], Mizanur Khondoker[1], Anne-Marie Minihane[1], Matthew G. Pontifex[1], Michael Müller[1], Simon McArthur[2], Gwenaelle Le Gall[1†] and David Vauzour[1†]*

[1] Norwich Medical School, Faculty of Medicine and Health Sciences, University of East Anglia, Norwich NR4 7TJ, United Kingdom.

[2] Institute of Dentistry, Faculty of Medicine & Dentistry, Queen Mary University of London, Blizard Institute, London E1 2AT, United Kingdom

† These authors share senior authorship

* To whom correspondence should be addressed: Dr David Vauzour; email: D.Vauzour@uea.ac.uk

Abstract

A key component of disease prevention is the identification of at-risk individuals. Microbial dysbiosis and microbe-derived metabolites (MDM) can influence the central nervous system, but their role in disease progression and as prognostic indicators is unknown. To identify preclinical factors associated with Alzheimer's disease (AD), we compared gut microbiome and metabolome profiles of cognitively healthy subjects, subjective cognitive impairment (SCI) participants and mild cognitive impairment (MCI) participants (n=50 per group, matched for age, BMI and sex), targeting metabolites previously associated with cognitive health (TMAO, bile acids, tryptophan, *p*-cresol and their derivatives). 16S rRNA bacterial microbiome sequencing and targeted LC-MS/MS were employed for faecal microbiome speciation and serum MDM quantification. Microbiome beta diversity differed between healthy controls and SCI participants. Multiple linear regression modelling highlighted five serum metabolites (indoxyl sulfate, choline, 5-hydroxyindole acetic acid, indole-3-propionic acid (IPA) and kynurenic acid) significantly altered in preclinical AD. Neuroprotective metabolites, including choline, 5-hydroxyindole acetic acid and IPA, exhibited lower concentrations in SCI and MCI in comparison to controls, while the cytotoxic metabolite indoxyl sulfate had higher levels. A Random Forest algorithm with multiclass classification confirmed and extended our results, identifying six metabolites (indoxyl sulfate, choline, 5-hydroxyindole acetic acid, IPA, kynurenic acid, kynurenine) as predictors of early cognitive decline, with an area under the curve of 0.74. In summary, a combined statistical and machine learning approach identified MDM as a novel composite risk factor for the early identification of future dementia risk.

Keywords: microbiome, indoxyl sulfate, indole-3-propionic acid, 5-hydroxyindole acetic acid, choline, kynurenic acid, mild cognitive impairment, subjective cognitive impairment.

1. Introduction

Currently, an estimated 55.2 million people suffer from dementia worldwide, of which Alzheimer's disease (AD) is the main form [1]. In the absence of an effective strategy to slow or prevent disease progression, dementia incidence is expected to increase to 152.8 million by 2050. By the time AD is typically diagnosed, substantial neuronal loss will have occurred across multiple brain regions. Identifying molecular precursors and biological risk factors at preclinical disease stages would enable earlier detection and the targeting of regular monitoring and mitigating interventions while prevention is viable.

The contribution of lifestyle factors to cognitive decline and dementia is well documented [2,3]. Diet in particular has emerged as a key influencer of brain health and AD development, in part by modulating communication along the microbiota-gut-brain axis. This axis forms a bidirectional communication system comprising neuronal, endocrine, immune and metabolic signalling mechanisms linking the gut and the central nervous system (CNS) [4]. Gut microbes regulate this communication via the breakdown of dietary compounds into bioactive metabolites. Such microbe-derived metabolites (MDM) subsequently modulate pathways affecting the CNS both directly, by crossing the blood-brain barrier and indirectly, via modulation of peripheral organ function or vagus nerve stimulation [5]. In the prodromal stages of AD, for example mild cognitive impairment (MCI), the microbiota-gut-brain axis becomes dysregulated (i.e., dysbiosis), a change associated with pathological processes such as neuroinflammation and neural injury, and thought to contribute to accelerating cognitive decline [6–8]. However, the mechanism(s) underlying these changes, and the role of MDM in this process remains unknown.

Several examples of MDM have been linked to cognitive health [5], including trimethylamine N-oxide (TMAO) [9–12], bile acids (BAs) [13–15], tryptophan [16–19], *p*-cresol and its derivatives [20,21]. Notably, these same MDM have been further linked to pathological processes known to be associated with AD [21–28], including neuroinflammation, synaptic damage and blood-brain barrier disruption, but whether changes in these MDM are drivers or correlates of disease processes requires comprehensive investigation.

Targeted metabolomics presents a powerful tool to comprehensively assess changes in the endogenous metabolome. Here, we present a targeted metabolomics approach employing liquid chromatography-tandem mass spectrometry (LC-MS/MS) to quantify TMAO, BAs, tryptophan and *p*-cresol metabolite profiles in the serum of healthy controls and participants in early cognitive decline. Early cognitive decline comprises individuals with subjective cognitive impairment (SCI) and mild cognitive impairment (MCI), the preclinical stages of AD progression. This study presents, for the first time, the prognostic value of key metabolites in combination and represents one of only a few

studies characterising metabolic perturbations in the early stages of cognitive decline, including participants undergoing the earliest preclinical stage of AD, SCI.

2. Materials and methods

2.1 Study samples

Human serum samples from the baseline measurements of two previously conducted clinical studies were used: (1) the impact of Cranberries On the Microbiome and Brain in healthy Ageing sTudy (COMBAT; NCT03679533) and (2) the Cognitive Ageing, Nutrition and Neurogenesis (CANN; NCT02525198) study. The COMBAT study recruited 60 adults, aged 50-80 years, with no subjective memory complaints as assessed by the Cognitive Change Index (CCI) questionnaire [29]. The CANN study recruited 259 participants, aged ≥ 50 years, with subjective cognitive impairment (SCI) or mild cognitive impairment (MCI) based on criteria developed by the National Institute of Aging-Alzheimers Association, with no indication of clinical dementia [30]. Cognitively healthy adults were selected from the COMBAT study as a control group, with all groups (controls, SCI and MCI, n=50 per group) matched for age, BMI and sex as these are key variables known to affect microbiome composition [31,32]. Participants with chronic fatigue syndrome, liver disease, diabetes mellitus, or gall bladder abnormalities were excluded.

Cognitive health was assessed using a variety of cognitive tests in both the COMBAT and CANN study. However, only the Trail Making Test (assessing visual processing speed, scanning, mental flexibility, as well as executive function) and the digit span test (assessing verbal short-term and working memory) were used across the COMBAT and CANN studies enabling comparisons. Participants also completed a validated, semi-quantitative Scottish Collaborative Group (SCG) food frequency questionnaire (version 6.6) to assess background diet [33]. Biochemical analyses of blood glucose, liver function (bilirubin, albumin, aspartate aminotransferase (AST), alanine aminotransferase (ALT) and AST/ALT ratio), kidney function (creatinine) and serum lipid concentrations (total-, LDL-, HDL-cholesterol and triglyceride) were conducted in all participants. The protocols were approved by the UK National Research Ethics Service (NRES) Committee, (Study ID: 14/EE/0189) for CANN and by the University of East Anglia's Faculty of Medicine and Health Sciences Ethical Review Committee (Reference: 201819–039) and the UK Health Research Authority (IRAS number: 237251) for COMBAT. The participants provided written informed consent to participate.

2.2 Microbiome Profiling

Microbiome analysis was performed by 16S rRNA sequencing as previously reported [34]. In brief, DNA extraction was performed from approximately 50 mg of faecal content using the QIAamp PowerFecal Pro DNA Kit (Qiagen, Manchester, UK) as per the manufacturer's instructions. DNA quantity was evaluated using a Nanodrop 2000 Spectrophotometer (Fisher Scientific, UK). Quality assessment by agarose gel electrophoresis distinguished the DNA integrity, purity, fragment size and concentration. Illumina NovaSeq 6000 PE250 was used to amplify the V3–V4 hypervariable region. Sequence analysis was carried out using Uparse software (Uparse v7.0.1001) [35], incorporating all the effective tags. Sequences sharing a similarity of ≥97% were grouped into the same Operational Taxonomic Unit (OTU). A representative sequence for each OTU was screened for further annotation. A representative OTU sequence was further analysed using the SSUrRNA database of SILVA Database 138 [36]. OTU abundance data were normalised using a standard sequence number corresponding to the sample with the least sequences. Alpha diversity was assessed using both Chao1 and Shannon H diversity indices, whilst beta diversity was assessed using Bray–Curtis dissimilarity. Statistical significance was determined by Kruskal–Wallis or Permutational Multivariate Analysis of Variance (PERMANOVA). Comparisons at the phylum and genus level were made using classical univariate analysis using Kruskal–Wallis combined with a false discovery rate (FDR) approach used to correct for multiple testing. P-values below 0.05 were considered statistically significant.

2.2 Metabolite Profiling

Serum samples were diluted with methanol at a ratio of 1:10 (v/v) and placed on dry ice for 10 min. Samples were then centrifuged (5 min, 16,000x g at room temp), supernatants filtered using a 0.45 µM PTFE syringe filter and evaporated to dryness using a Savant™ SpeedVac™ High-Capacity Concentrator (Thermofisher, UK). Dried samples were resuspended in either 50 µL of methanol with the addition of 15 µL of lithocholic acid-d4 and cholic acid-d4 at 50 µg/mL for the detection of bile acids, 50 µL water with TMA-d9 N-oxide, TMA 13C 15N hydrochloride at 50 µg/mL for the detection of TMAO/TMA/choline or 50 µL water with 15 µL of L-methionine-3, 3, 4, 4 d4 and p-toluenesulfonic acid at 50 µg/mL for the detection of tryptophan and *p*-cresol metabolites respectively. All internal standards were supplied by Thermofisher, UK. Samples were analysed using the Waters Acquity UPLC system and Xevo TQ-S Cronos mass spectrometer with MassLynx 4.1 software. See supplementary methods for full details.

2.3 Statistical Analyses

Significant associations of metabolites with cognitive status were identified using multiple linear regression analysis. Covariates known to affect metabolome or microbiome composition, including age, BMI, diet and markers of kidney function (creatinine) and liver function (AST/ALT ratio), were

included in the model [37–41]. Sex was not included as a covariate as all groups had an equal proportion of males and females. Diet was assessed using a validated, semi-quantitative Scottish Collaborative Group (SCG) food frequency questionnaire (version 6.6) [33]. Participants' dietary components were grouped (kcal, proteins, fats, carbohydrates, water, alcohol, vitamins and minerals) and analysed using hierarchical clustering via Ward's linkage method to assemble individuals with similar dietary patterns (Supplementary Figure S1). This clustered participants into low, moderate and high intake of dietary components and was added to the model as a categorical variable, with participants with a moderate intake used as a reference group. Age, BMI, creatinine and AST/ALT ratio were added to the model as continuous variables. Finally, cognitive status (i.e., control, SCI and MCI) was added to the model as a categorical variable. Metabolite concentrations outside ± 2 standard deviations from the mean were excluded as outliers. The assumptions for multiple linear regression analysis including the existence of a linear relationship among the outcome and predictor variable, normality and homoscedasticity were assessed (Supplementary Figure S2). The model tested for significant associations between metabolite and cognitive status, adjusting for the included covariates. All multiple linear regression analyses were performed in R (v3.6.3; R Foundation: A Language and Environment for Statistical Computing).

2.4 Machine Learning

A Random Forest (RF) machine learning algorithm was implemented to assess whether metabolites could be predictive of preclinical AD. The RF model was constructed using 100 decision trees and 6 random variables considered at each split. The number of variables considered per split corresponds to the square root of the total number of attributes in the data [42] (as 32 variables were considered, this resulted in ~6 random variables per split). To create a composite panel to predict preclinical AD, metabolites were ranked according to the mean decrease Gini. This highlights the loss in model performance when permuting the predictor values, and can provide more robust results than mean decrease accuracy [43]. The metabolites with the highest mean decrease Gini score producing the highest AUC values were retained in the model. To compare our model, Naive Bayes and AdaBoost machine learning models were also constructed [44,45]. AdaBoost predictions were made by using a weighted average of weak classifiers. Our model contained 50 estimators, a learning rate of 1.00 and a SAMME classification algorithm which updated the base estimator's weights with classification results. The Naive Bayes method was applied based on applying Bayes' theorem with the "naive" assumption of conditional independence between every pair of features given the value of the target variable. The dataset for multi-class classification was randomly divided into training and testing, with 75% of the samples allocated to the training set and 25% to the testing set. Models were assessed by the average area under the receiver operator curve (AUC) (plotting the false positive rate against

the true positive rate) over all classes (macro-average) as an indication of model performance. All machine learning models were built in Python (Python Software Foundation. Python Language Reference, version 3.8).

3. Results

3.1 Study population characteristics

A total of 150 individuals were included in the study of which 50 (33.3%) were cognitively healthy, 50 (33.3%) presented with SCI and 50 (33.3%) with MCI. The mean ± SD age of all participants was 65.5 ± 5.7 years, with a mean level of education of 14.6 ± 3.5 years and 54% female (Table 1). Cognitive groups were matched for age, BMI and sex (p= 0.99). Participants in both the COMBAT and CANN study undertook several cognitive assessments at their baseline visit [29,30]. Significant differences were found in the Trail Making Test B, digit span backward test and digit span total score between groups ($p<0.05$). There was a marginal difference between the three groups in the Trail Making Test A (p= 0.09) and no significant difference occurred in the digit span forwards test (p= 0.21). The prevalence of the *APO □4* was lower in controls (18%) compared to SCI (26%) and MCI (38%) participants.

Albumin, bilirubin and fasting glucose ($p<0.01$) differed according to cognitive status. Interestingly, both albumin and bilirubin were highest in controls and lowest in SCI participants. Although participants diagnosed with diabetes mellitus were excluded, fasting glucose increased over preclinical AD, with the lowest concentrations in control individuals and the highest in MCI (Table 1).

3.2 Gut microbiome and metabolome shifts in preclinical AD

Alpha diversity was measured using the Chao1 (p= 0.21) and Shannon H (p= 0.70) indices with no significant difference amongst groups (Figure 1A-B). Conversely, beta diversity, as measured by Bray-Curtis dissimilarity, was significantly different (PERMANOVA F-value= 1.35, p= 0.02) (Figure 1C). Pairwise analysis suggested the shift was primarily driven by the differences between the control and SCI groups (FDR q= 0.03), rather than between SCI and MCI (FDR q= 0.38) or MCI and control (FDR q= 0.15) (Figure 1D). The PLS-DA plot suggested similar patterns in participants' metabolomic profiles, with control separating from SCI and MCI (Figure 2A). The extent of this separation can be seen through the heatmap displaying shifts in metabolite concentrations between groups and clustering SCI and MCI together (Figure 2B). The similarity between microbiome and metabolomic profiles was confirmed by conducting a Procrustes analysis to evaluate the congruence of the two

datasets. The analysis revealed strong similarity between the metabolome and microbiome results in control and SCI (R= 0.21, p= 0.03), SCI and MCI (R= 0.27, p= 0.002) and MCI and control (R= 0.26, p= 0.002) groups (Supplementary Figure S3).

3.3 Serum metabolites significantly associated with early cognitive decline in an adjusted multivariable model

Multiple linear regression analysis adjusted for liver function (AST/ALT ratio), kidney function (creatinine), age, BMI and background diet identified five metabolites significantly associated with early cognitive decline, including choline, 5-hydroxyindole acetic acid, indole-3- propionic acid (IPA), indoxyl sulfate and kynurenic acid (Table 1). Indoxyl sulfate, choline and 5-hydroxyindole acetic acid were associated with both SCI and MCI ($p<0.05$). Kynurenic acid was significantly associated with SCI (β= 0.007, 95% CI: <0.001, 0.014, p= 0.037) but not MCI (β= 0.001, 95% CI: -0.006, 0.007, p= 0.874). On the other hand, IPA was significantly associated with MCI (β= -0.558, 95% CI: -0.910, -0.206, p= 0.002), but not SCI (β= -0.181, 95% CI: -0.536, 0.174, p= 0.316). Neuroprotective metabolites, including choline, 5-hydroxyindole acetic acid and indole propionic acid [46–48], exhibited lower concentrations in SCI and MCI participants in comparison to controls, while metabolites linked to cytotoxicity, including indoxyl sulfate, showed increasing levels [49]. Kynurenic acid, a typically neuroprotective metabolite [50], was higher in SCI and MCI in comparison to controls. Group means for all metabolite concentrations are given in Supplementary Table S1.

3.4 Machine learning models to identify risk factors predictive of preclinical AD

All 32 serum metabolites were initially evaluated as possible predictors of preclinical AD. RF achieved the highest classification AUC of 0.65, with AdaBoost and Naïve Bayes attaining 0.58 and 0.63 respectively (Supplementary Table S2). Using the mean decrease Gini, the importance of each metabolite was assessed (Figure S4). Six metabolites (5-hydroxyindole acetic acid, indole-3-propionic acid, choline, indoxyl sulfate, kynurenic acid and kynurenine) produced the highest AUC of 0.74 using the RF classification algorithm (Figure 4). In comparison, Naive Bayes achieved an AUC of 0.72 and AdaBoost attained 0.68. The RF ROC curve indicated the model's predictive performance was highest for controls (AUC= 0.79), followed by MCI (AUC= 0.76) and SCI (AUC= 0.64). As such, we investigated whether the model performance would be improved by predicting only healthy ageing and MCI. Using the six serum metabolites from controls and MCI participants, the RF model showed improved predictive performance (AUC= 0.84) (Supplementary Table S2). AdaBoost and Naive Bayes also demonstrated increased performance, attaining AUC of 0.87 and 0.90 respectively.

3.5 Serum dietary and MDM and gut microbiome modulation in preclinical AD are significantly linked

Having identified a shift in both the microbiome and metabolome profiles, a Spearman correlation examined possible connections between the two datasets. Significantly modulated metabolites and microbiome genera between the three groups ($p<0.05$) were correlated, revealing bacterial–metabolite interactions. Control and SCI participants displayed a negative relationship between 5-hydroxyindole acetic acid and *Lachnoclostridium* (R= -0.29, p=0.004), indoxyl sulfate and *Turicibacter* (R= -0.21, p= 0.038), and choline and *Lachnoclostridium* (R= -0.36, p= 0.0002) and *Ruminococcus gnavus* group (R= -0.28, p= 0.005) (Figure 3A). Choline and *UCG-009* (R= 0.316, p= 0.002), anthranilic acid and *Lactonifactor* (R= 0.24, p= 0.019) and *Holdemanella* (R= 0.24, p= 0.02), indoxyl sulfate and *Holdemania* (R= 0.318, p= 0.001) and *Lactonifactor* (R= 0.33, p=0.0008) had a positive correlation. Between SCI and MCI participants, only IPA and *Lachnospiraceae ND3007* group had a positive correlation (R= 0.26, p= 0.011) (Figure 3B).

As indoxyl sulfate, choline, 5-hydroxyindole acetic acid and IPA were identified as risk factors of cognitive decline, genera correlated with these metabolites were investigated further to identify potential associations with their production. *Holdemania* and *Lachnoclostridium* genera were higher in SCI participants in comparison to controls but lower in SCI relative to MCI (Figure 3C-D). On the other hand, *Turicibacter* were higher in both SCI and MCI participants in comparison to controls (Figure 3E). The abundance of *UCG-009* was significantly lower in both SCI and MCI participants in comparison to controls (Figure 3F). *Lactonifactor* and *Ruminococcus gnavus* were both higher in SCI in comparison to controls (Figure G-H). Finally, *Lachnospiraceae ND3007* was not different between control and SCI but was lower in SCI compared to MCI (Figure 3I).

4. Discussion

Identification of robust, inexpensive and non-invasive markers of cognitive status and its trajectory is currently an unmet medical need in AD research, with circulating gut-derived metabolites presenting a promising area. Metabolic alterations contain rich systemic information on the underlying physiology that connects the periphery to the CNS, likely affecting numerous pathways simultaneously. Thus, the simultaneous detection of numerous perturbed metabolites can provide a powerful detection tool. However, studies investigating composite markers are lacking.

16s rRNA sequencing indicated that significant shifts in gut microbiome composition occur during preclinical AD, commencing as early as SCI, suggesting changes may already be apparent when memory complaints first appear, aligning with previous studies [51,52]. As cognitive decline progresses from SCI to MCI, gut microbiome modulation appears to be less significant. Circulatory metabolites

also reflect this pattern, clustering SCI and MCI participants independently from the healthy controls. Individuals with SCI are likely at a higher risk of cognitive decline progression compared to those who are cognitively healthy [53], which may lead to greater alterations in biological markers, including the gut microbiome and its metabolites. Procrustes analysis showed significant congruence of the microbiome and metabolome datasets, suggesting the two are interlinked and provides the potential for MDM predictors to be detected early in disease progression. Indeed, gut microbiome composition can account for up to 58% of the variation of circulatory metabolites communicating along the microbiota-gut-brain axis [54].

Targeted metabolomics quantifies metabolites with extremely high sensitivity and accuracy, providing an advantage over the relative responses yielded by untargeted approaches. RF and multiple linear regression models both revealed indoxyl sulfate, choline, 5-hydroxyindole acetic acid, IPA and kynurenic acid as key early indicators of cognitive decline, with RF presenting an AUC predictive performance of 0.74, strongly supporting a significant link between metabolic perturbations associated with the gut microbiome and preclinical AD progression. Previous studies have predominantly concentrated on binary classification approaches, primarily utilising MRI and PET imaging modalities, to investigate AD progression [55–57]. However, in clinical practice, multiclass classification of blood samples of patients with SCI, MCI and healthy controls could provide a useful approach. Tong and colleagues attained a similar predictive performance (AUC= 0.729) using RF and nonlinear graph fusion of multiple modalities (regional MRI volumes, voxel-based FDG-PET signal intensities, CSF biomarker measures and genetic information) to classify control, MCI and AD participants [58]. AUC increased to 0.84 when predicting healthy ageing and MCI, likely due to the difficulty of diagnosing a patient undergoing SCI. Indeed, Purser and colleagues found no relationship between memory complaints and the progression of cognitive impairment over 10 years in individuals 65 years and over [59]. However, others dispute this result [60]. Adjusting our statistical analysis for confounding variables that heavily influence the host, such as age, BMI, kidney function, liver function and background diet, improves analysis robustness and sensitivity. Adjusting for background diet becomes particularly vital when examining MDM as the diet can both modulate gut microbiome composition and provide a variety of bioactive precursor compounds; a factor which is often overlooked in metabolomic analyses [61]. Nevertheless, our results highlight the use of profiling select circulatory MDM to identify higher-risk individuals of cognitive decline.

Of the five metabolites highlighted by both machine learning and multiple linear regression, all except choline are produced from tryptophan metabolism, indicating notable alterations in tryptophan metabolism may occur in preclinical AD progression. Tryptophan metabolism has previously been well-linked to AD [18]. Indeed, we find lower neuroprotective tryptophan-derived metabolites, including IPA and 5-hydroxyindole acetic acid, as cognitive decline progresses from controls to SCI and MCI. IPA is produced in the gut by the microbial conversion of tryptophan via the indole pathway and has

previously been investigated as a possible treatment for AD [62] due to its potent antioxidant effect against Aβ 1-42 *in vitro* [63] and its ability to prevent aggregation and deposition of Aβ monomers [64]. IPA is anti-inflammatory, reducing the concentration of the proinflammatory TNF-α in activated microglia [65], lowering the expression of chemokine (CC Motif) ligand 2 (CCL2) and nitric oxide synthase 2 (NOS2) in interferon-beta (IFN-β) activated murine astrocytes [66] and preventing increases in cytokines in LPS-induced human primary astrocytes [67], and has previously been identified as a predictor of AD progression [68]. 5-hydroxyindole acetic acid is often used as a surrogate marker for serotonin due to serotonin's rapid degradation. As such, our findings indicate lower peripheral serotonin breakdown as early cognitive decline progresses. Approximately 95% of all serotonin is localised in peripheral compartments where it is involved in the modulation of the enteric nervous system (ENS) development and neurogenesis, gut motility, secretion, inflammation, and epithelial development, suggesting these processes may be disrupted in early cognitive decline [69]. Indeed, MCI and AD patients have often been reported to suffer from gastrointestinal symptoms [70] and ENS dysregulation in AD has previously been described [71]. Decreased concentrations of 5-hydroxyindole acetic acid also suggest a shift in tryptophan metabolism towards the kynurenine pathway, reducing the availability of tryptophan for serotonin synthesis. This is supported by higher serum kynurenine concentrations in SCI and MCI participants in comparison to controls (Supplementary Table S1) and has previously been found in AD participants, linked to poor memory, executive function and global cognition [72]. The kynurenine pathway is activated by an inflammatory stimulus, promoting indoleamine 2,3-dioxygenase, the rate-limiting enzyme that initiates the kynurenine pathway. Increased inflammation is a common feature of AD and as such may play a role in modulating tryptophan catabolites.

Both indoxyl sulfate and kynurenic acid concentrations were increased as cognitive decline progressed, even after adjusting for measures of liver and kidney function. As a uremic toxin, indoxyl sulfate can disrupt neuronal efflux transport systems, promote the production of free radicals, inflammation, endothelial cell dysfunction and disturb the circadian rhythm involved in clearing renal and CNS toxins [73,74], likely contributing to cognitive decline. Serum levels of indoxyl sulfate, as well as albumin, have previously been identified as predictive of cognitive impairment in participants with end-stage renal disease [75]. End-stage renal disease patients have also been reported to have an increased abundance of the gut bacteria *Holdemania*, in line with our results, suggesting his genera may be underlying the changes between control and SCI [76]. Rodent studies show increased kynurenic acid concentrations can impair cognitive function, including spatial working memory, and broad monitoring deficits [77,78]. However, data regarding this relationship in human studies is inconsistent [79,80]. Kynurenic acid can play a protective role against the cytotoxic product of the kynurenine pathway, quinolinic acid, by acting as an NMDA antagonist for both glycine and glutamate modulatory sites [81]. However, abnormal accumulation has previously been found to induce

glutamatergic hypofunction and subsequently disrupt cognitive function [82]. In AD, increased blood concentrations of kynurenic acid have been hypothesised to relate to neuroinflammatory processes and may be produced as a protective response to neuronal damage [83].

Choline is required for numerous biological functions in the body [84], notably including hallmark AD-associated processes such as acetylcholine synthesis [85]. As choline readily crosses the blood-brain barrier, peripheral concentrations typically mirror concentrations in the brain [86], thus lower concentrations in early cognitive decline may indicate decreased central acetylcholine production. Acetylcholine is intricately connected to neural networks regulating memory, and a reduction in this system is closely associated with learning and memory deficits in AD [87]. *Lachnoclostridium* and *Lactonifactor* were inversely correlated with choline levels, suggesting changes in these genera may modulate blood concentrations. Indeed, previous research has found *L. saccharolyticum* WM1, a representative strain of *Lachnoclostridium*, to be an efficient converter of choline to TMA *in vitro*, transforming at a rate near 100% [88]. This metabolic process *in vivo* also elevated serum TMAO levels [88], which is supported by our results displaying a 1.6-fold higher TMAO in MCI compared to controls. It is likely that increases in *Lachnoclostridium* abundance may increase the metabolism of choline to TMAO, decreasing its concentration in circulation.

Our study has major strengths including simultaneously targeting some of the top microbial and metabolic metabolites associated with cognitive decline, whilst matching our participants and adjusting our analysis for key factors known to influence the metabolome (age, BMI, sex, liver function, kidney function and background diet), factors rarely accounted for in marker studies. Furthermore, our study highlights key microbiota underlying these metabolic changes, as well as investigating participants from the earliest stage of decline (SCI) and validating our results through machine learning and adjusted statistical approaches. However, some limitations should be stressed. Despite our study adjusting results for key covariates, host metabolome profiles are influenced by a plethora of additional largely environmental and biological factors. Thus, although our findings suggest relationships between the variables, we cannot infer causal relationships from this analysis alone. Moreover, participants' background diet was adjusted for using data collected by food frequency questionnaires, which can be prone to measurement error and may introduce inaccuracies due to recall bias and self-reporting issues. Furthermore, like all studies utilising machine learning models, the larger the dataset, the more robust the predictive performance. With the current dataset including 150 individuals and 32 metabolites, we achieved significant predictive performance. However, our findings will require external validation in larger independent cohorts to improve the model.

Pathophysiological progression of AD is apparent up to 20 years prior to clinical symptom onset, making it vital for prevention research to focus on uncovering novel preclinical risk factors. Scalable

markers that enable the early detection of at-risk persons could permit the targeting of lifestyle interventions to lessen future risk and uncover novel mechanisms underpinning dementia. Our findings present new insights into the preclinical progression of cognitive decline and dementia. We signify a major role for the gut in connection to the brain through the modulation of key MDM. Furthermore, we lend strength to the hypothesis that individuals with higher risks of cognitive decline can be identified via a targeted metabolomic approach in the preceding stages of AD.

Data availability

The 16S rRNA gene sequence data have been deposited in the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/) under accession number PRJNA1109848. Other data that support the findings of this study are available from the corresponding authors upon reasonable request.

Authors Contributions

D.V. conceptualised the project. A.M.M. provided the human blood and faecal samples. G.L.G. designed the metabolomic analysis, which E.C. and G.L.G. then performed. E.C. conducted the 16S rRNA microbiome sequencing. E.C. and S.S. performed the machine learning analysis. E.C. and M.K. conducted the statistical analysis. E.C. wrote the paper. D.V., S.S., M.K., A.M.M., M.G.P., M.M., S.M. G.L.G. contributed key discussions in interpreting results and edited the paper. All authors approved the final version of the paper.

Figure Legends

Figure 1: Microbiome beta diversity is significantly altered in early cognitive decline. Alpha diversity measured by Chao1 (A) and Shannon H (B) index. (C) Beta diversity as measured by Bray-curtis; p-value generated from PERMANOVA. (D) Pairwise comparisons of the beta diversity analysis.

Figure 2: Metabolic shift occurs in early cognitive decline. (A) Partial least squares-discriminant analysis (PLS-DA) plot of the metabolomic profiles. (B) Heatmap displaying changes in concentrations of metabolites between the groups, with hierarchical clustering.

Figure 3: Serum metabolome and gut microbiome profiles are linked. Spearman rank correlation analysis between metabolite and microbiome genera that are significantly modulated in early cognitive decline (A) between control and SCI and (B) between SCI and MCI. (C-H) Abundance counts of microbiome genera correlated with our metabolites of interest (indoxyl sulfate, choline and 5-hydroxyindole acetic acid) between control and SCI participants. (I) Abundance count of microbiome genera correlated with metabolite of interest (indole-3-propionic acid) between SCI and MCI, *=$p<0.05$, **=$p<0.01$.

Figure 4: Six circulatory metabolites are predictive of preclinical AD. Receiving Operating Characteristic (ROC) curve illustrating the performance of the Random Forest model for classifying controls, SCI and MCI participants with average area under the curve (AUC) of the multilevel classifier.

5. References

1. World Health Organization. Global status report on the public health response to dementia. *Security Research Hub Reports* (2021).
2. Flicker, L. Modifiable Lifestyle Risk Factors for Alzheimer's Disease. *Journal of Alzheimer's Disease* **20**, 803–811 (2010).
3. Livingston, G. *et al.* Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet* **396**, 413–446 (2020).
4. Chakrabarti, A. *et al.* The microbiota-gut-brain axis: pathways to better brain health. Perspectives on what we know, what we need to investigate and how to put knowledge into practice. *Cell Mol Life Sci* **79**, 80 (2022).
5. Connell, E. *et al.* Microbial-derived metabolites as a risk factor of age-related cognitive decline and dementia. *Molecular Neurodegeneration* **17**, 43 (2022).
6. Li, B. *et al.* Mild cognitive impairment has similar alterations as Alzheimer's disease in gut microbiota. *Alzheimer's & Dementia* **15**, 1357–1366 (2019).
7. Liu, P. *et al.* Altered microbiomes distinguish Alzheimer's disease from amnestic mild cognitive impairment and health in a Chinese cohort. *Brain, Behavior, and Immunity* **80**, 633–643 (2019).
8. Nagpal, R. *et al.* Gut mycobiome and its interaction with diet, gut bacteria and alzheimer's disease markers in subjects with mild cognitive impairment: A pilot study. *EBioMedicine* **59**, 102950 (2020).
9. Brunt, V. E. *et al.* The gut microbiome–derived metabolite trimethylamine N-oxide modulates neuroinflammation and cognitive function with aging. *GeroScience* **43**, 377–394 (2021).
10. Li, D. *et al.* Trimethylamine-N-oxide promotes brain aging and cognitive impairment in mice. *Aging Cell* **17**, e12768 (2018).
11. Zhuang, Z. *et al.* Causal relationships between gut metabolites and Alzheimer's disease: a bidirectional Mendelian randomization study. *Neurobiology of Aging* **100**, 119.e15-119.e18 (2021).

12. Hoyles, L. *et al.* Regulation of blood-brain barrier integrity by microbiome-associated methylamines and cognition by trimethylamine N-oxide. *Microbiome* **9**, 235 (2021).
13. Baloni, P. *et al.* Metabolic Network Analysis Reveals Altered Bile Acid Synthesis and Metabolism in Alzheimer's Disease. *Cell Rep Med* **1**, 100138 (2020).
14. MahmoudianDehkordi, S. *et al.* Altered Bile Acid Profile Associates with Cognitive Impairment in Alzheimer's Disease – An Emerging Role for Gut Microbiome. *Alzheimers Dement* **15**, 76–92 (2019).
15. Nho, K. *et al.* Altered Bile Acid Profile in Mild Cognitive Impairment and Alzheimer's Disease: Relationship to Neuroimaging and CSF Biomarkers. *Alzheimers Dement* **15**, 232–244 (2019).
16. Hestad, K., Alexander, J., Rootwelt, H. & Aaseth, J. O. The Role of Tryptophan Dysmetabolism and Quinolinic Acid in Depressive and Neurodegenerative Diseases. *Biomolecules* **12**, 998 (2022).
17. Schwarcz, R., Bruno, J. P., Muchowski, P. J. & Wu, H.-Q. Kynurenines in the mammalian brain: when physiology meets pathology. *Nat Rev Neurosci* **13**, 465–477 (2012).
18. Weaver, D., Gupta, M., Meek, A., Wang, Y. & Wu, F. Alzheimer's Disease as a Disorder of Tryptophan Metabolism (2745). *Neurology* **94**, (2020).
19. Whiley, L. *et al.* Metabolic phenotyping reveals a reduction in the bioavailability of serotonin and kynurenine pathway metabolites in both the urine and serum of individuals living with Alzheimer's disease. *Alzheimer's Research & Therapy* **13**, 20 (2021).
20. Sankowski, B. *et al.* Higher cerebrospinal fluid to plasma ratio of p-cresol sulfate and indoxyl sulfate in patients with Parkinson's disease. *Clinica Chimica Acta* **501**, 165–173 (2020).
21. Shah, S. N. *et al.* Cerebrovascular damage caused by the gut microbe-derived uraemic toxin p-cresol sulfate is prevented by blockade of the epidermal growth factor receptor. 2022.11.12.516113 Preprint at https://doi.org/10.1101/2022.11.12.516113 (2022).
22. Smith, M. J., Crowther, R. A. & Goedert, M. The natural osmolyte trimethylamine N-oxide (TMAO) restores the ability of mutant tau to promote microtubule assembly. *FEBS Lett* **484**, 265–270 (2000).

23. Jena, P. K. *et al.* Dysregulated bile acid synthesis and dysbiosis are implicated in Western diet-induced systemic inflammation, microglial activation, and reduced neuroplasticity. *FASEB J* **32**, 2866–2877 (2018).
24. Tran, C. D. *et al.* Gut permeability, its interaction with gut microflora and effects on metabolic health are mediated by the lymphatics system, liver and bile acid. *Future Microbiology* **10**, 1339–1353 (2015).
25. Azevedo, M. L. V. *et al.* p-Cresyl sulfate affects the oxidative burst, phagocytosis process, and antigen presentation of monocyte-derived macrophages. *Toxicol Lett* **263**, 1–5 (2016).
26. Edamatsu, T., Fujieda, A. & Itoh, Y. Phenyl sulfate, indoxyl sulfate and p-cresyl sulfate decrease glutathione level to render cells vulnerable to oxidative stress in renal tubular cells. *PLoS One* **13**, e0193342 (2018).
27. Sun, C.-Y., Cheng, M.-L., Pan, H.-C., Lee, J.-H. & Lee, C.-C. Protein-bound uremic toxins impaired mitochondrial dynamics and functions. *Oncotarget* **8**, 77722–77733 (2017).
28. Tang, W.-H. *et al.* Protein-bounded uremic toxin p-cresylsulfate induces vascular permeability alternations. *Histochem Cell Biol* **149**, 607–617 (2018).
29. Flanagan, E. *et al.* Chronic Consumption of Cranberries (Vaccinium macrocarpon) for 12 Weeks Improves Episodic Memory and Regional Brain Perfusion in Healthy Older Adults: A Randomised, Placebo-Controlled, Parallel-Groups Feasibility Study. *Frontiers in Nutrition* **9**, (2022).
30. Irvine, M. A. *et al.* The Cognitive Ageing, Nutrition and Neurogenesis (CANN) trial: Design and progress. *Alzheimers Dement (N Y)* **4**, 591–601 (2018).
31. Haro, C. *et al.* Intestinal Microbiota Is Influenced by Gender and Body Mass Index. *PLOS ONE* **11**, e0154090 (2016).
32. Zhang, X. *et al.* Sex- and age-related trajectories of the adult human gut microbiota shared across populations of different ethnicities. *Nat Aging* **1**, 87–100 (2021).
33. Hollis, J. L. *et al.* Assessing the relative validity of the Scottish Collaborative Group FFQ for measuring dietary intake in adults. *Public Health Nutrition* **20**, 449–455 (2017).

34. Pontifex, M. G. *et al.* Saffron extract (Safr'Inside™) improves anxiety related behaviour in a mouse model of low-grade inflammation through the modulation of the microbiota and gut derived metabolites. *Food Funct.* **13**, 12219–12233 (2022).
35. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73**, 5261–5267 (2007).
36. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**, D590-596 (2013).
37. Gowda, S. *et al.* Markers of renal function tests. *N Am J Med Sci* **2**, 170–173 (2010).
38. Williams, A. L. & Hoofnagle, J. H. Ratio of serum aspartate to alanine aminotransferase in chronic hepatitis. Relationship to cirrhosis. *Gastroenterology* **95**, 734–739 (1988).
39. Giannini, E. *et al.* Progressive liver functional impairment is associated with an increase in AST/ALT ratio. *Dig Dis Sci* **44**, 1249–1253 (1999).
40. Tavassol, Z. H. *et al.* Alteration in Gut Microbiota Composition of Older Adults Is Associated with Obesity and Its Indices: A Systematic Review. *The Journal of nutrition, health and aging* **27**, 817–823 (2023).
41. Leeming, E. R., Johnson, A. J., Spector, T. D. & Le Roy, C. I. Effect of Diet on the Gut Microbiota: Rethinking Intervention Duration. *Nutrients* **11**, 2862 (2019).
42. Liaw, A. & Wiener, M. Classification and Regression by RandomForest. *Forest* **23**, (2001).
43. Calle, M. L. & Urrea, V. Letter to the editor: Stability of Random Forest importance measures. *Brief Bioinform* **12**, 86–89 (2011).
44. Wang, R. AdaBoost for Feature Selection, Classification and Its Relation with SVM, A Review. *Physics Procedia* **25**, 800–807 (2012).
45. Rish, I. An Empirical Study of the Naïve Bayes Classifier. *IJCAI 2001 Work Empir Methods Artif Intell* **3**, (2001).
46. Hwang, I. K. *et al.* Indole-3-propionic acid attenuates neuronal damage and oxidative stress in the ischemic hippocampus. *J Neurosci Res* **87**, 2126–2137 (2009).

47. Blusztajn, J. K., Slack, B. E. & Mellott, T. J. Neuroprotective Actions of Dietary Choline. *Nutrients* **9**, 815 (2017).

48. Klein, C. *et al.* 5-HIAA induces neprilysin to ameliorate pathophysiology and symptoms in a mouse model for Alzheimer's disease. *Acta Neuropathologica Communications* **6**, 136 (2018).

49. Leong, S. C. & Sirich, T. L. Indoxyl Sulfate—Review of Toxicity and Therapeutic Strategies. *Toxins* **8**, 358 (2016).

50. Ostapiuk, A. & Urbanska, E. M. Kynurenic acid in neurodegenerative disorders—unique neuroprotection or double edged sword? *CNS Neurosci Ther* **28**, 19–35 (2021).

51. Sheng, C. *et al.* Altered Gut Microbiota in Adults with Subjective Cognitive Decline: The SILCODE Study. *J Alzheimers Dis* **82**, 513–526 (2021).

52. Chen, G., Zhou, X., Zhu, Y., Shi, W. & Kong, L. Gut microbiome characteristics in subjective cognitive decline, mild cognitive impairment and Alzheimer's disease: a systematic review and meta-analysis. *Eur J Neurol* **30**, 3568–3580 (2023).

53. Reisberg, B., Shulman, M. B., Torossian, C., Leng, L. & Zhu, W. Outcome over seven years of healthy adults with and without subjective cognitive impairment. *Alzheimers Dement* **6**, 10.1016/j.jalz.2009.10.002 (2010).

54. Dekkers, K. F. *et al.* An online atlas of human plasma metabolite signatures of gut microbiome composition. *Nat Commun* **13**, 5370 (2022).

55. Simfukwe, C., Youn, Y. C., Kim, M.-J., Paik, J. & Han, S.-H. CNN for a Regression Machine Learning Algorithm for Predicting Cognitive Impairment Using qEEG. *Neuropsychiatric Disease and Treatment* **19**, 851–863 (2023).

56. Subramanyam Rallabandi, V. P. & Seetharaman, K. Deep learning-based classification of healthy aging controls, mild cognitive impairment and Alzheimer's disease using fusion of MRI-PET imaging. *Biomedical Signal Processing and Control* **80**, 104312 (2023).

57. Zubrikhina, M. O. *et al.* Machine learning approaches to mild cognitive impairment detection based on structural MRI data and morphometric features. *Cognitive Systems Research* **78**, 87–95 (2023).

58. Tong, T., Gray, K., Gao, Q., Chen, L. & Rueckert, D. Multi-modal classification of Alzheimer's disease using nonlinear graph fusion. *Pattern Recognition* **63**, 171–181 (2017).

59. Purser, J. L., Fillenbaum, G. G. & Wallace, R. B. Memory complaint is not necessary for diagnosis of mild cognitive impairment and does not predict 10-year trajectories of functional disability, word recall, or short portable mental status questionnaire limitations. *J Am Geriatr Soc* **54**, 335–338 (2006).

60. Geerlings, M. I., Jonker, C., Bouter, L. M., Adèr, H. J. & Schmand, B. Association between memory complaints and incident Alzheimer's disease in elderly people with normal baseline cognition. *Am J Psychiatry* **156**, 531–537 (1999).

61. Playdon, M. C. *et al.* Comparing metabolite profiles of habitual diet in serum and urine. *Am J Clin Nutr* **104**, 776–789 (2016).

62. Bendheim, P. E. *et al.* Development of indole-3-propionic acid (OXIGON) for Alzheimer's disease. *J Mol Neurosci* **19**, 213–217 (2002).

63. Chyan, Y. J. *et al.* Potent neuroprotective properties against the Alzheimer beta-amyloid by an endogenous melatonin-related indole structure, indole-3-propionic acid. *J Biol Chem* **274**, 21937–21942 (1999).

64. Cheng, X. & van Breemen, R. B. Mass spectrometry-based screening for inhibitors of beta-amyloid protein aggregation. *Anal Chem* **77**, 7012–7015 (2005).

65. Kim, C.-S., Jung, S., Hwang, G.-S. & Shin, D.-M. Gut microbiota indole-3-propionic acid mediates neuroprotective effect of probiotic consumption in healthy elderly: A randomized, double-blind, placebo-controlled, multicenter trial and in vitro study. *Clinical Nutrition* **42**, 1025–1033 (2023).

66. Rothhammer, V. *et al.* Type I interferons and microbial metabolites of tryptophan modulate astrocyte activity and central nervous system inflammation via the aryl hydrocarbon receptor. *Nat Med* **22**, 586–597 (2016).

67. Garcez, M. L., Tan, V. X., Heng, B. & Guillemin, G. J. Sodium Butyrate and Indole-3-propionic Acid Prevent the Increase of Cytokines and Kynurenine Levels in LPS-induced Human Primary Astrocytes. *Int J Tryptophan Res* **13**, 1178646920978404 (2020).

68. Gao, J. *et al.* The Association of Fried Meat Consumption With the Gut Microbiota and Fecal Metabolites and Its Impact on Glucose Homoeostasis, Intestinal Endotoxin Levels, and Systemic Inflammation: A Randomized Controlled-Feeding Trial. *Diabetes Care* **44**, 1970–1979 (2021).
69. Terry, N. & Margolis, K. G. Serotonergic Mechanisms Regulating the GI Tract: Experimental Evidence and Therapeutic Relevance. *Handb Exp Pharmacol* **239**, 319–342 (2017).
70. Rao, M. & Gershon, M. D. The bowel and beyond: the enteric nervous system in neurological disorders. *Nat Rev Gastroenterol Hepatol* **13**, 517–528 (2016).
71. Chalazonitis, A. & Rao, M. Enteric nervous system manifestations of neurodegenerative disease. *Brain Res* **1693**, 207–213 (2018).
72. Willette, A. A. *et al.* Inflammation, negative affect, and amyloid burden in Alzheimer's disease: Insights from the kynurenine pathway. *Brain, Behavior, and Immunity* **95**, 216–225 (2021).
73. Franco, Á. de O., Starosta, R. T. & Roriz-Cruz, M. The specific impact of uremic toxins upon cognitive domains: a review. *J Bras Nefrol* **41**, 103–111 (2019).
74. Iwata, K. *et al.* Involvement of indoxyl sulfate in renal and central nervous system toxicities during cisplatin-induced acute renal failure. *Pharm Res* **24**, 662–671 (2007).
75. Hou, Y.-C., Lu, K.-C., Huang, C.-L., Lin, Y.-F. & Chen, R.-M. Serum albumin and indoxyl sulfate were predictive of cognitive impairment via amyloid beta and tauopathy, respectively, in end-stage renal disease patients. *Alzheimer's & Dementia* **17**, e053921 (2021).
76. Zheng, L. J. *et al.* Gut dysbiosis-influence on amygdala-based functional activity in patients with end stage renal disease: a preliminary study. *Brain Imaging Behav* **14**, 2731–2744 (2020).
77. Hahn, B., Reneski, C. H., Pocivavsek, A. & Schwarcz, R. Prenatal kynurenine treatment in rats causes schizophrenia-like broad monitoring deficits in adulthood. *Psychopharmacology (Berl)* **235**, 651–661 (2018).
78. Pocivavsek, A. *et al.* Fluctuations in endogenous kynurenic acid control hippocampal glutamate and memory. *Neuropsychopharmacology* **36**, 2357–2367 (2011).
79. Fazio, F. *et al.* Xanthurenic Acid Activates mGlu2/3 Metabotropic Glutamate Receptors and is a Potential Trait Marker for Schizophrenia. *Sci Rep* **5**, 17799 (2015).

80. Chiappelli, J. *et al.* Stress-induced increase in kynurenic acid as a potential biomarker for patients with schizophrenia and distress intolerance. *JAMA Psychiatry* **71**, 761–768 (2014).
81. Stone, T. W. & Darlington, L. G. The kynurenine pathway as a therapeutic target in cognitive and neurodegenerative disorders. *Br J Pharmacol* **169**, 1211–1227 (2013).
82. Fujigaki, H., Yamamoto, Y. & Saito, K. L-Tryptophan-kynurenine pathway enzymes are therapeutic target for neuropsychiatric diseases: Focus on cell type differences. *Neuropharmacology* **112**, 264–274 (2017).
83. Marrugo-Ramírez, J., Rodríguez-Núñez, M., Marco, M.-P., Mir, M. & Samitier, J. Kynurenic Acid Electrochemical Immunosensor: Blood-Based Diagnosis of Alzheimer's Disease. *Biosensors* **11**, 20 (2021).
84. Zeisel, S. H. & da Costa, K.-A. Choline: An Essential Nutrient for Public Health. *Nutr Rev* **67**, 615–623 (2009).
85. Dave, N. *et al.* Dietary choline intake is necessary to prevent systems-wide organ pathology and reduce Alzheimer's disease hallmarks. *Aging Cell* **22**, e13775 (2023).
86. Wurtman, R. J., Cansev, M. & Ulus, I. H. Choline and Its Products Acetylcholine and Phosphatidylcholine. in *Handbook of Neurochemistry and Molecular Neurobiology: Neural Lipids* (eds. Lajtha, A., Tettamanti, G. & Goracci, G.) 443–501 (Springer US, Boston, MA, 2009). doi:10.1007/978-0-387-30378-9_18.
87. Poly, C. *et al.* The relation of dietary choline to cognitive performance and white-matter hyperintensity in the Framingham Offspring Cohort1234. *Am J Clin Nutr* **94**, 1584–1591 (2011).
88. Cai, Y.-Y. *et al.* Integrated metagenomics identifies a crucial role for trimethylamine-producing Lachnoclostridium in promoting atherosclerosis. *npj Biofilms Microbiomes* **8**, 1–12 (2022).

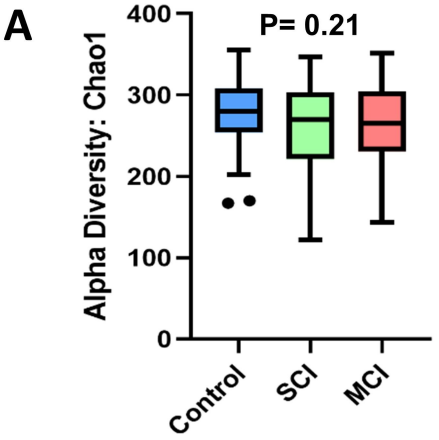
Table 1: **Baseline characteristics of the participants**. Mean (SD). P-value calculated using one-way ANOVA. Significant values at p<0.05 are in bold. SCI: subjective cognitive impairment, MCI: mild cognitive impairment, TMT: Trail Making Test A or B. Bold p-values represent p<0.05.

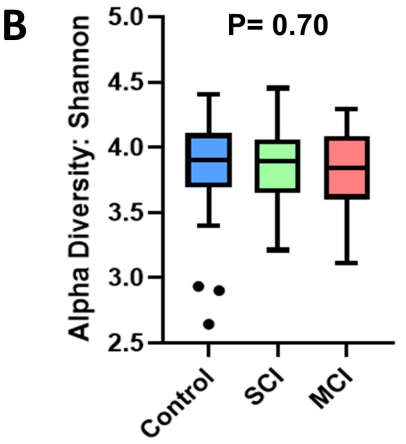
	Control (N=50)	**SCI (N=50)**	**MCI (N=50)**	**P-value**
Sex, M/ F (%F)	23/27 (54)	23/27 (54)	23/27 (54)	-
Age (years)	65.6 (5.3)	65.5 (6.1)	65.5 (5.8)	0.999
BMI (kg/m^2)	25.1 (3.1)	25.0 (2.9)	25.0 (2.8)	0.993
Education (years)	14.4 (2.6)	14.6 (4.0)	14.6 (3.9)	0.968
% *APOE4*	18	26	38	0.079
Cognitive tests				
TMT A	30.7 (6.2)	29.3 (8.1)	33.3 (12.1)	0.088
TMAT B	66.4 (20.4)	62.3 (16.5)	74.9 (27.3)	**0.015**
Digit Span Forwards	11.1 (2.2)	11.2 (1.8)	10.5 (2.6)	0.211
Digit Span Backwards	7.7 (2.0)	7.2 (1.8)	6.5 (2.1)	**0.011**
Digit Span Total	18.8 (3.8)	18.4 (3.0)	17.0 (4.2)	**0.039**
Biochemistry				
Creatinine (µmol/L)	73.90 (13.5)	72.5 (12.3)	73.5 (14.1)	0.871
Albumin (g/L)	40.4 (2.4)	30.0 (2.4)	39.4 (2.3)	**<0.001**
Bilirubin (µmol/L)	12.8 (4.7)	8.9 (5.1)	9.1 (4.4)	**<0.001**
AST (µL)	21.7 (3.9)	20.6 (5.9)	24.0 (13.1)	0.141
ALT (µL)	16.9 (5.4)	16.7 (9.0)	18.3 (11.3)	0.621
AST/ALT	1.4 (0.3)	1.4 (0.4)	1.4 (0.4)	0.548
Fasting Glucose (mmol/L)	4.8 (0.4)	5.0 (0.5)	5.3 (1.0)	**<0.001**
Triglyceride (mmol/L)	1.1 (0.5)	1.1 (0.4)	1.2 (0.5)	0.368
Cholesterol (mmol/L)	5.6 (1.1)	5.2 (1.1)	5.2 (1.0)	0.174
HDL Cholesterol (mmol/L)	1.6 (0.4)	1.5 (0.5)	1.5 (0.4)	0.297
LDL Cholesterol (mmol/L)	3.4 (0.1)	3.2 (0.9)	3.1 (0.8)	0.182

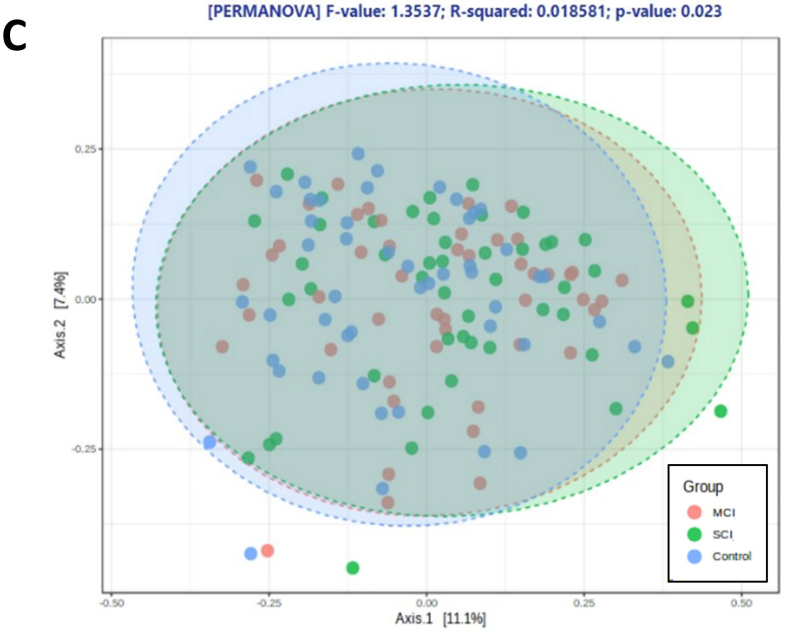
Table 2: Multiple linear regression model (adjusted for age, BMI, liver function (AST/ALT ratio), kidney function (creatinine) and diet) showing metabolites significantly associated with early cognitive decline. Diet was analysed using hierarchical clustering, 'Ward' method, to group individuals with similar dietary patterns. This grouped participants into three dietary groups (low, moderate and high intake of dietary components (Kcal, carbohydrates, fats, protein, water, alcohol, minerals, vitamins). Healthy controls and diet group 2 (moderate intake) were used as reference groups in the model. Bold p-values represent $p<0.05$.

Explanatory Variable	**Metabolite**																			
	Indoxyl Sulfate				**Choline**				**5-Hydroxyindole Acetic Acid**				**Indole Propionic Acid**				**Kynurenic Acid**			
	Beta	*P-value*	*95% CI*		*Beta*	*P-value*	*95% CI*		*Beta*	*P-value*	*95% CI*		*Beta*	*P-value*	*95% CI*		*Beta*	*P-value*	*95% CI*	
			Low	*High*			*Low*	*High*			*Low*	*High*			*Low*	*High*			*Low*	*High*
Constant	-2.288	0.482	-8.705	4.128	29.690	**0.011**	6.938	52.443	0.068	**0.020**	0.011	0.125	3.094	**0.009**	0.790	5.397	0.017	0.448	-0.027	0.062
Age	0.057	0.115	-0.014	0.128	0.100	0.434	-0.152	0.351	0.001	0.058	0.000	0.001	-0.007	0.568	-0.033	0.018	<0.001	0.314	-0.001	<0.001
BMI	0.064	0.378	-0.079	0.207	-0.328	0.202	-0.835	0.178	-0.001	0.063	-0.003	0.000	-0.026	0.308	-0.077	0.025	0.001	**0.039**	<0.001	0.002
Creatinine	0.036	**0.028**	0.004	0.067	0.105	0.063	-0.006	0.217	0.000	0.493	0.000	0.000	-0.004	0.452	-0.016	0.007	0.001	**<0.001**	<0.001	0.001
AST/ALT	-1.437	**0.011**	-2.541	-0.332	0.812	0.681	-3.082	4.706	-0.002	0.710	-0.012	0.008	0.213	0.289	-0.182	0.607	-0.009	**0.023**	-0.017	-0.001
Diet Group 1	0.354	0.445	-0.559	1.267	1.247	0.445	-1.972	4.466	-0.002	0.690	-0.010	0.006	0.204	0.221	-0.124	0.532	-0.003	0.408	-0.009	0.004
Diet Group 3	0.130	0.829	-1.052	1.311	-0.706	0.740	-4.897	3.486	-0.001	0.826	-0.012	0.009	-0.128	0.568	-0.570	0.314	-0.003	0.450	-0.011	0.005
SCI	1.650	**0.001**	0.674	2.625	-5.635	**0.002**	-9.084	-2.187	-0.011	**0.012**	-0.020	-0.002	-0.181	0.316	-0.536	0.174	0.007	**0.037**	<0.001	0.014
MCI	1.308	**0.008**	0.342	2.274	-5.217	**0.003**	-8.645	-1.789	-0.015	**0.001**	-0.024	-0.007	-0.558	**0.002**	-0.910	-0.206	0.001	0.874	-0.006	0.007
Overall change in early cognitive decline	Increased				Decreased				Decreased				Decreased				Increased			

638

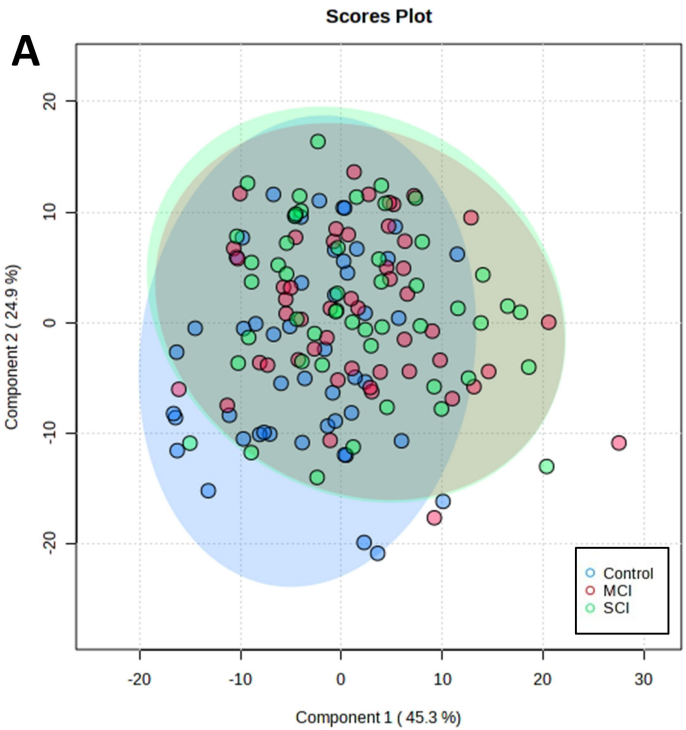


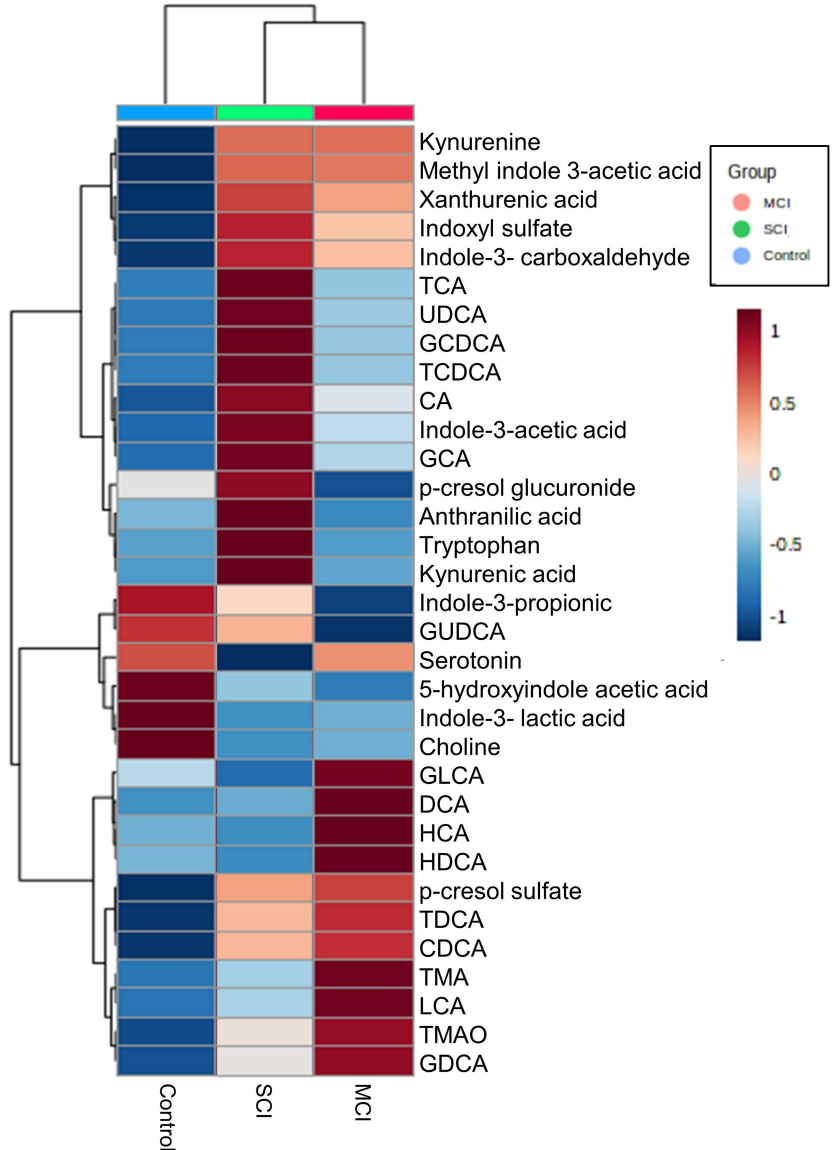


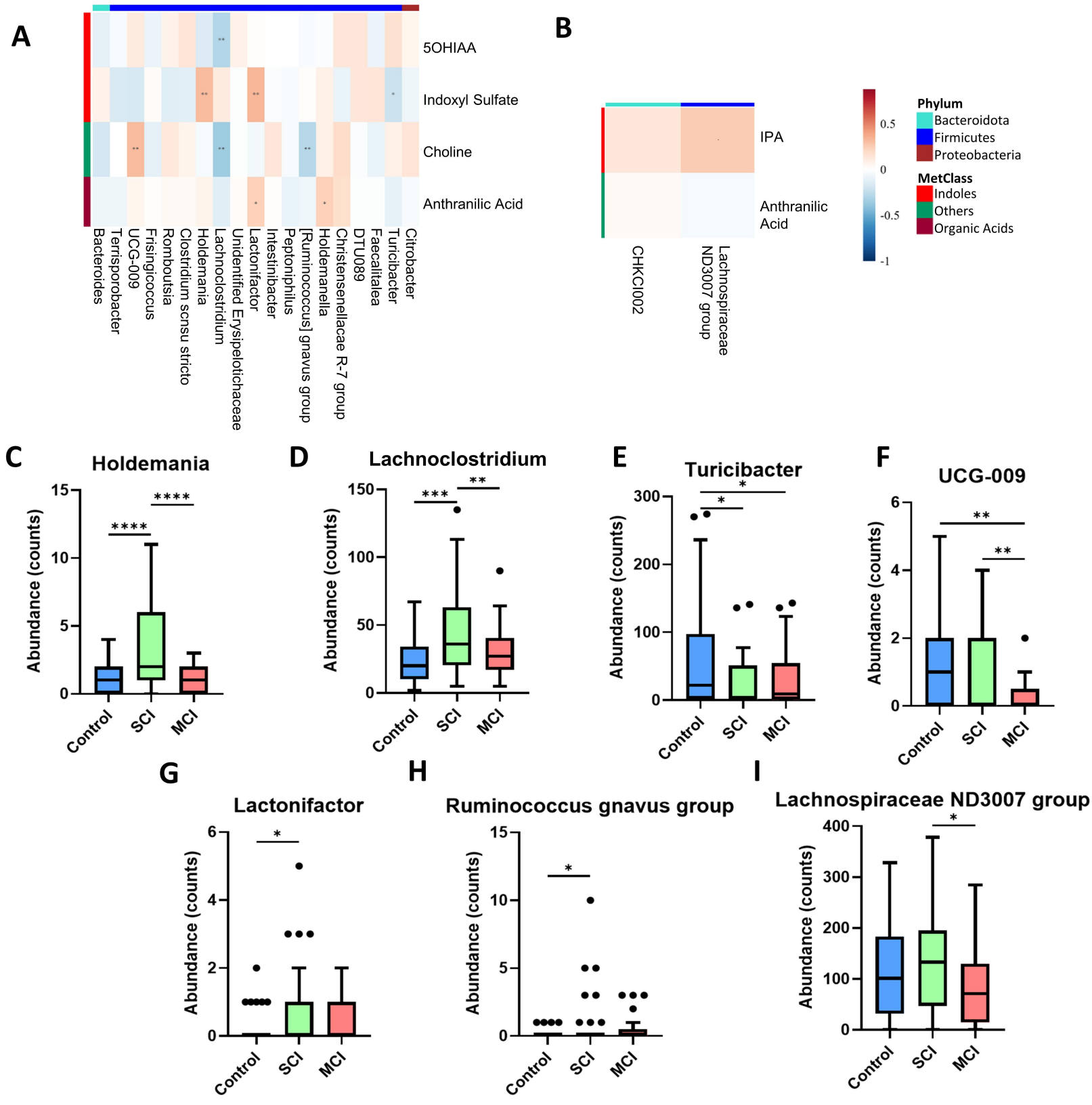


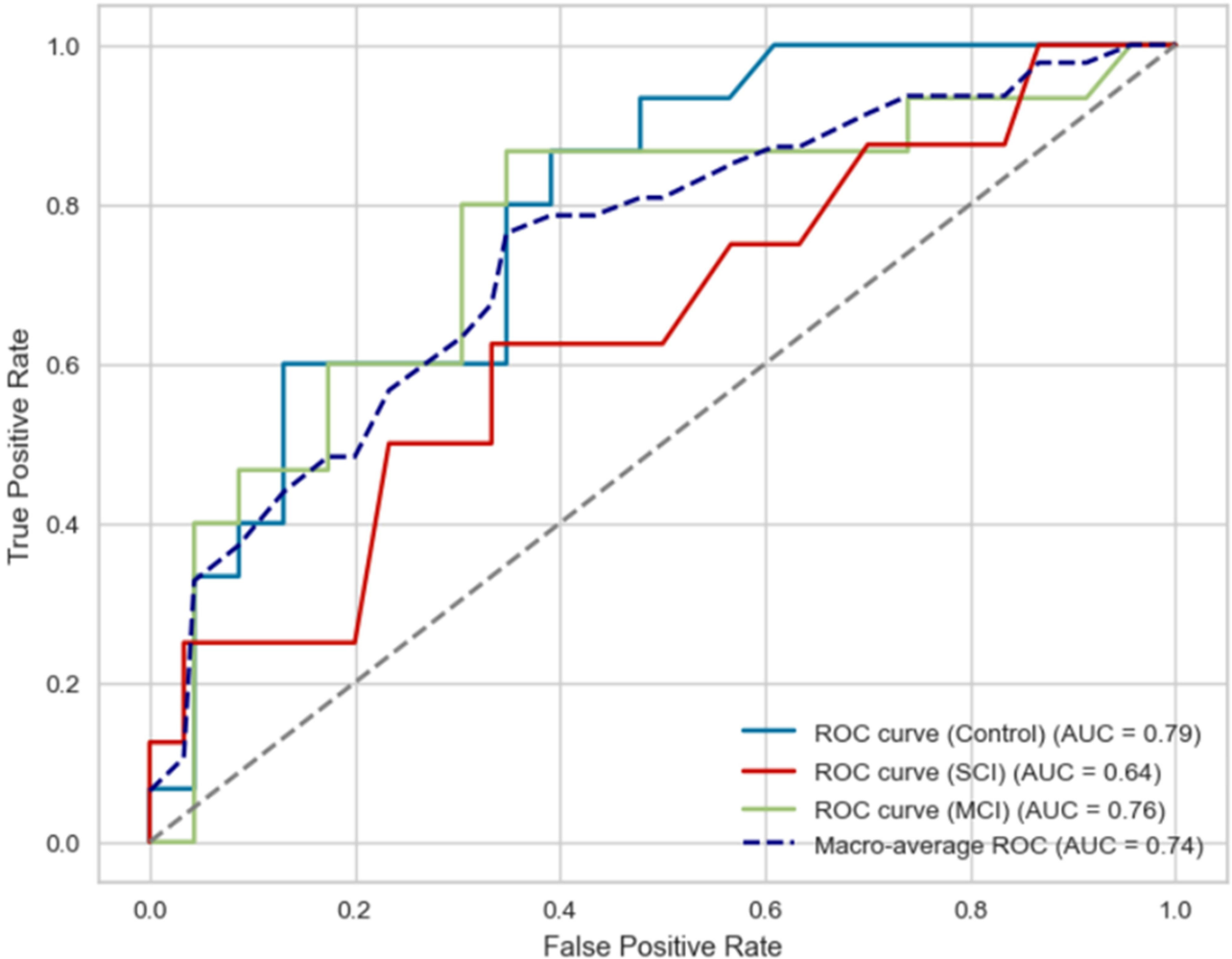
D

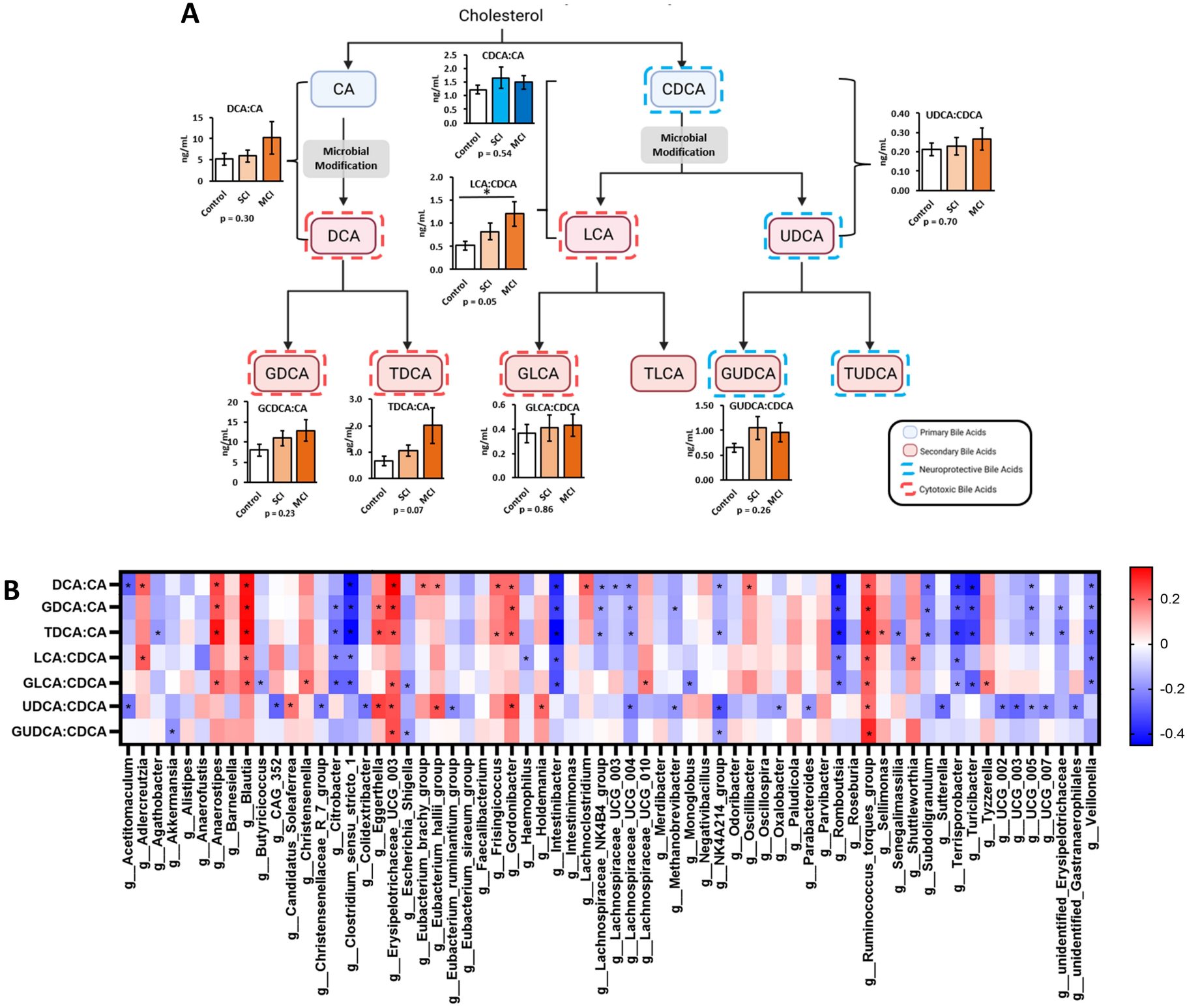
Pair	F-value	R-Squared	P-value	FDR
Control vs SCI	1.77	0.02	0.01 *	0.03 *
SCI vs MCI	1.02	0.01	0.38	0.38
MCI vs Control	1.26	0.01	0.10	0.15


A
Scores Plot
Component 2 (24.9 %)
Component 1 (45.3 %)
20
10
0
-10
-20
-20
-10
0
10
20
30
Control
MCI
SCI


Kynurenine
Methyl indole 3-acetic acid
Xanthurenic acid
Indoxyl sulfate
Indole-3- carboxaldehyde
TCA
UDCA
GCDCA
TCDCA
CA
Indole-3-acetic acid
GCA
p-cresol glucuronide
Anthranilic acid
Tryptophan
Kynurenic acid
Indole-3-propionic
GUDCA
Serotonin
5-hydroxyindole acetic acid
Indole-3- lactic acid
Choline
GLCA
DCA
HCA
HDCA
p-cresol sulfate
TDCA
CDCA
TMA
LCA
TMAO
GDCA
Control
SCI
MCI
Group
MCI
SCI
Control
1
0.5
0
-0.5
-1


A
5OHIAA
Indoxyl Sulfate
Choline
Anthranilic Acid
Bacteroides
Terrisporobacter
UCG-009
Frisingicoccus
Romboutsia
Clostridium sensu stricto
Holdemania
Lachnoclostridium
Unidentified Erysipelotichaceae
Lactonifactor
Intestinibacter
Peptoniphilus
[Ruminococcus] gnavus group
Holdemanella
Christensenellaceae R-7 group
DTU089
Faecalitalea
Turicibacter
Citrobacter
B
IPA
Anthranilic Acid
CHKCI002
Lachnospiraceae ND3007 group
Phylum
Bacteroidota
Firmicutes
Proteobacteria
MetClass
Indoles
Others
Organic Acids
C
Holdemania
D
Lachnoclostridium
E
Turicibacter
F
UCG-009
G
Lactonifactor
H
Ruminococcus gnavus group
I
Lachnospiraceae ND3007 group
Abundance (counts)
Control
SCI
MCI


1.0
0.8
0.6
0.4
0.2
0.0
True Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
False Positive Rate
ROC curve (Control) (AUC = 0.79)
ROC curve (SCI) (AUC = 0.64)
ROC curve (MCI) (AUC = 0.76)
Macro-average ROC (AUC = 0.74)


A
Cholesterol
CA
CDCA
Microbial Modification
Microbial Modification
DCA
LCA
UDCA
GDCA
TDCA
GLCA
TLCA
GUDCA
TUDCA
DCA:CA
p = 0.30
CDCA:CA
p = 0.54
LCA:CDCA
p = 0.05
UDCA:CDCA
p = 0.70
GCDCA:CA
p = 0.23
TDCA:CA
p = 0.07
GLCA:CDCA
p = 0.86
GUDCA:CDCA
p = 0.26
ng/mL
Control
SCI
MCI
Primary Bile Acids
Secondary Bile Acids
Neuroprotective Bile Acids
Cytotoxic Bile Acids
B
DCA:CA
GDCA:CA
TDCA:CA
LCA:CDCA
GLCA:CDCA
UDCA:CDCA
GUDCA:CDCA
g__Acetitomaculum
g__Adlercreutzia
g__Agathobacter
g__Akkermansia
g__Alistipes
g__Anaerofustis
g__Anaerostipes
g__Barnesiella
g__Blautia
g__Butyricicoccus
g__CAG_352
g__Candidatus_Soleaferrea
g__Christensenella
g__Christensenellaceae_R_7_group
g__Citrobacter
g__Clostridium_sensu_stricto_1
g__Colidextribacter
g__Eggerthella
g__Erysipelotrichaceae_UCG_003
g__Escherichia_Shigella
g__Eubacterium_brachy_group
g__Eubacterium_hallii_group
g__Eubacterium_ruminantium_group
g__Eubacterium_siraeum_group
g__Faecalibacterium
g__Frisingicoccus
g__Gordonibacter
g__Haemophilus
g__Holdemania
g__Intestinibacter
g__Intestinimonas
g__Lachnoclostridium
g__Lachnospiraceae_NK4B4_group
g__Lachnospiraceae_UCG_003
g__Lachnospiraceae_UCG_004
g__Lachnospiraceae_UCG_010
g__Merdibacter
g__Methanobrevibacter
g__Monoglobus
g__Negativibacillus
g__NK4A214_group
g__Odoribacter
g__Oscillibacter
g__Oscillospira
g__Oxalobacter
g__Paludicola
g__Parabacteroides
g__Parvibacter
g__Romboutsia
g__Roseburia
g__Ruminococcus_torques_group
g__Sellimonas
g__Senegalimassilia
g__Shuttleworthia
g__Subdoligranulum
g__Sutterella
g__Terrisporobacter
g__Turicibacter
g__Tyzzerella
g__UCG_002
g__UCG_003
g__UCG_005
g__UCG_007
g__unidentified_Erysipelotrichaceae
g__unidentified_Gastranaerophilales
g__Veillonella
0.2
0
-0.2
-0.4