

METHODS SUPPLEMENT

Participant Information:

All participants were deemed to be at risk for AD based on their systemic metabolic dysfunction and cognitive dysfunction or subjective memory complaints at study onset (see **Table 1** for full participant characteristics). Criteria for metabolic dysfunction was a hemoglobin A1c of level 5.7-6.4%, which corresponds to the American Diabetes Association pre-diabetes category according to their 2016 guidelines (1). Subjective memory complaints (SMC) were diagnosed using the Alzheimer's Disease Neuroimaging Initiative (ADNI) criteria and mild cognitive impairment (MCI) was diagnosed by expert physicians and neuropsychologists using NIA-AA guidelines for clinical diagnoses without reference to biomarker status (2, 3). Study exclusion criteria included prior diagnosis of neurological or neurodegenerative illness (except MCI), major psychiatric disorder (although well-controlled depression was allowed), prior stroke, current use of diabetes and lipid lowering medications, or medications with known effects on the central nervous system (i.e. anti-seizure medications, anti-psychotics, opioids, etc.).

The protocol was approved by the Wake Forest Institutional Review Board (ClinicalTrials.gov Identifier: NCT02984540), and written informed consent was obtained from all participants and/or their study partners. Participants were medically supervised by clinicians, with safety monitoring overseen by the Wake Forest Institutional Data and Safety Monitoring Committee.

Procedure:

The primary pilot trial was a randomized crossover design in which participants consumed either a Modified Mediterranean-Ketogenic Diet (MMKD) or the control American Heart Association Diet (AHAD) for 6 weeks, followed by a 6-week washout, after which the second diet was consumed for 6 weeks (**Figure S1**). During the washout period participants were instructed to resume their pre-study diet and specifically not continue the protocols from their first diet. Randomization was performed with a random number generator so that the number of MCI and SMC participants in each dietary group would be roughly equal. Baseline cognitive, imaging, blood, stool, and CSF measures were obtained prior to diet randomization (4).

Diet Intervention and Education:

The experimental Modified Mediterranean Ketogenic Diet (MMKD) was a very low carbohydrate diet aimed at inducing ketosis. Notably, modified versions of the KD have increasingly been utilized in cases of medically intractable epilepsy given its similar efficacy and increased tolerability for patients (5). These modified KDs generally incorporate slightly higher amounts of carbohydrates from whole food sources, yet are still very low carbohydrate and reliably induce ketosis. The control diet was adapted from the low-fat American Heart Association Diet (AHAD, 5-6). The diets were customized to each participant's baseline caloric needs to maintain their current weight throughout the course of the study. Participants were also asked to keep their exercise and physical activity level stable throughout the study.

The proportions of carbohydrates and fat were the main variables manipulated between the two diets. The target macronutrient composition (expressed as % of total calories) was 5-10% carbohydrate, 60-65% fat, and 30% protein for the MMKD and 55-65% carbohydrate, 15-20% fat, and 20-30% protein for the AHAD. Participants on the MMKD were asked to keep their daily carbohydrate consumption below 20g each day throughout the 6-week intervention, while higher fat foods (preferably those low in saturated fats) could be consumed liberally. They were encouraged to eat fish, lean meats, leafy green vegetables, and nuts and discouraged from consuming artificially sweetened beverages or products marketed as “low-carbohydrate” during the intervention. Participants on the AHAD were encouraged to limit their amount of fat intake to 40g or less each day, and to eat plentiful fruits, vegetables, and fiber-laden carbohydrates.

A registered dietitian developed daily meal plans for each participant based upon their food preferences and caloric needs as determined by a pre-study 3-day food diary, body composition, and activity level. Participants had weekly diet education visits (either in-person or by phone) starting one week prior to the start of each diet and continuing throughout the remainder of the intervention. Participants maintained a daily food record that was reviewed at these visits.

Adherence to the dietary intervention was assessed by capillary ketone body (beta-hydroxybutyrate) measures that were collected at all major time points and during diet education visits using the Nova Max Plus® (<http://www.novacares.com/nova-max-plus/>) and with participant subjective report recorded by study dietitian. Prior evidence suggests that less frequent blood-based ketone body measures are equally accurate as daily urine ketone body test strips in assessment of ketosis (7).

For the most part, participants were required to supply their own food. They were given a food stipend of \$25 each week provided to offset the higher food costs associated with switching their diets. All participants were provided with a daily multivitamin supplement (Centrum® Silver®), and participants on the MMKD were supplied with 1L bottles of extra virgin olive oil during their Pre-Diet and Mid-Diet visits to promote consumption of healthy fats. Participants were not allowed to use resveratrol, CoQ10 (coenzyme Q10), curcumin, coconut oil or other medium chain triglyceride-containing supplements throughout the duration of the study.

Stool Collection

Stool samples were collected from participants at five study timepoints (**Figure S1**). Methods for collection were adapted from the Manual of Procedures for the Human Microbiome Project (NIH, Version Number 12.0). Participants received gloves, a sterile collection container, sterile scooping tool, and toilet hat in addition to instructions for proper specimen collection and storage (refrigeration or on ice) prior to transfer to study staff. Once received, samples were immediately aliquoted into four individual vials from each timepoint and stored at -80°C until further processing and analyses.

Metagenomic Sample Processing:

DNA was extracted from stool samples according to Earth Microbiome Project protocols (8) using the QIAGEN® MagAttract® PowerSoil® DNA KF Kit (384-sample). A total of 5 ng (or 3.5 µL maximum) genomic DNA were used in a 1:10 miniaturized Kapa HyperPlus protocol with a 15-cycle PCR amplification for shotgun metagenomic sequencing. Libraries were quantified with the PicoGreen dsDNA assay kit, and 50 ng (or 1 µL maximum) of each library was pooled. The pool was size selected for 300 to 700 bp and sequenced as a paired-end 150-cycle run on an Illumina HiSeq 4000 sequencer at the UCSD IGM Genomics Center (9).

Metagenomic Data Processing:

Shotgun sequencing data were uploaded to and processed by Qiita (10, Study ID 13662). Human reads were removed using minimap2 2.17 (11), while adapters, quality filtering, and trimming were performed using fastp 20.1 (12). Remaining reads were recruited to the Web of Life database (13) with Bowtie2 v2.3.0 (14) using the parameters from the SHOGUN pipeline (15), then processed into Operational Genomic Units with Woltka (16). The resulting feature table was used for downstream analysis.

Metabolomic Sample Processing:

To extract metabolites from fecal samples, frozen samples were thawed on ice for 30 minutes. Then, a solution of 50% methanol spiked with 1 µM sulfamethazine was added to each fecal sample (approximately 50-100 mg feces) at a volume ratio of 10 µL extraction solvent to 1 mg sample. Samples were homogenized at 25 Hz for 5 minutes on a tissue homogenizer, then centrifuged at maximum speed for 15 minutes at 4°C. A 200 µL aliquot of supernatant from each sample was transferred into a 96-well plate and vacuum concentrated to dryness via centrifugal lyophilization (Labconco Centriva). Once dried, the samples were stored at -80 °C until LC-MS was performed. Untargeted LC-MS was performed using a Vanquish liquid chromatography system (Thermo Scientific) paired with QExactive mass spectrometer. Samples were separated using a C18 column (Phenomenex Kinetex 1.7 µm C18 100 Å LC Column 50 x 2.1 mm). The mobile phase used was LC-MS grade water (phase A) and LC-MS grade acetonitrile (phase B), both containing 0.1% formic acid (Fisher Scientific, Optima LC-MS). The LC was programmed with a flow rate set to 0.5 mL/min. Samples were injected at 95%A:5%B, which was held for 1 minute, before ramping up to 100%B over 7 minutes, which was held for 0.5 minutes before returning to starting conditions. The orbitrap mass spectrometer was equipped with a HESI-II ESI probe. The mass spectrometer was programmed to use a data-dependent acquisition method that acquired MS full scan spectra, followed by MS/MS spectra of the top 5 most abundant ions. Precursor ions were fragmented once before being added to an exclusion list for 30 seconds. Data were collected in positive ion mode.

Metabolomics Data Processing:

Raw Q Exactive files were converted to .mzXML format using the ProteoWizard tool MSConvert (17) software, then these reformatted files were imported to MZmine (version 2.53; 18). We performed feature finding using the parameters recommended for Feature Based Molecular Networking (FBMN) in the Global Natural Product Social Molecular Networking (GNPS) ecosystem in its documentation (19, 20). Only peaks that had an MS2 scan were retained. We exported the feature quantification table (.csv) and MS/MS spectral summary (.mgf) from

MZmine into GNPS and performed FBMN using release 28.2 with the default parameters (19, 20)

Food-omics Data Processing:

Food counts were inferred from metabolomics data using the “Global FoodOmics” project (<http://www.globalfoodomics.org>) reference data set. This dataset contains 3,579 food and beverage samples contributed by community members, that were systematically annotated (21). Matching up metabolite spectra from stool samples to metabolites identified in reference food samples enables estimation of foods consumed without the need for food frequency questionnaires (22). Specifically, each “food count” corresponded to an individual consensus node in the molecular networking results that matched to a node found in reference food sample molecular networks. Consensus nodes were required to match the specific sample type, GFOP, and not match to any of the other experiment groups. Infrequent food types that occurred less often than water (which is considered blank) were removed to filter out sporadic random matches.

Dimensionality Reduction Analysis:

Initial dimensionality reduction analysis was performed by Robust Aitchison Principal Component Analysis (RPCA, 23) through the gemelli plugin to QIIME2 (24; <https://github.com/biocore/gemelli>). Dimensionality reduction that accounted for the repeated measures study design was performed with compositional tensor factorization (CTF, 25), also through the gemelli plugin for QIIME2. The biplots output from these methods were visualized with EMPeror (26; <https://github.com/biocore/emperor>). The statistical significance of results were evaluated with ANOVAs comparing beta-diversity distances to baseline (27). Multiple test correction was performed through Bonferroni correction.

Relative Abundance of Features Analysis:

We identified metagenomic, food-omics, and metabolomic features that are associated with diet and cognition by performing Bayesian inferential regression (<https://github.com/gibbsramen/BIRDMAn>). We utilized a Negative Binomial Linear Mixed Effect model to ensure that our statistics were in agreement with our study design; we modeled time, dietary sequence (whether MMKD intervention was first or second), cognitive status, and diet (whether a given individual was on MMKD or AHAD at a given time point) as fixed effects while subject identity was modeled as a random effect. Microbes, food features, and metabolites were ordered by the log ratio of their relative abundances in objectively normal cognition to mild cognitive impairment or MMKD to AHAD individuals, respectively (28) and the top and bottom ten features were examined more closely.

Multi-omics Analysis:

Microbe and metabolite co-occurrence probabilities were calculated with MMvec’s paired-omics function (29; <https://github.com/biocore/mmvec>). Heatmap visualizations were also made with MMvec. The metabolic capabilities of the microbes were evaluated with MetaCyc (30) on the Web of Life genomes (13). Any genomes containing BSH (E.C. 3.5.1.24) were considered BSH-containing/encoding.

Data Availability:

Metagenomic sequencing data are available in Qiita (10) under study ID 13662. Mass spectrometry data (.mzXML format) are available in MassIVE under ID MSV000087087. The classical molecular networking job is available in GNPS at the following link: <https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=90f1ba6e1e4d4d89b75b9017a0631983>. The feature-based molecular networking job is available in GNPS at the following link: <https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=16e7c66221ce4ae9ae6678ec276d8343>.

The code utilized for these analyses are available at https://github.com/ahdilmore/BEAM_MultiOmics.

SUPPLEMENTAL REFERENCES

1. American Diabetes Association. Standards of Medical Care in Diabetes-2016 Abridged for Primary Care Providers. Clin Diabetes 2016;34:3–21.
2. Risacher SL, Kim S, Nho K, Foroud T, Shen L, Petersen RC, et al. APOE effect on Alzheimer's disease biomarkers in older adults with significant memory concern. Alzheimers Dement 2015;11:1417–29.
3. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The Diagnosis of Mild Cognitive Impairment due to Alzheimer's Disease: Recommendations from the National Institute on Aging-Alzheimer's Association Workgroups on Diagnostic Guidelines for Alzheimer's Disease. FOCUS 2013;11:96–106. <https://doi.org/10.1176/appi.focus.11.1.96>.
4. Neth BJ, Mintz A, Whitlow C, Jung Y, Sai KS, Register TC, et al. Modified ketogenic diet is associated with improved cerebrospinal fluid biomarker profile, cerebral perfusion, and cerebral ketone body uptake in older adults at risk for Alzheimer's disease: a pilot study. Neurobiology of Aging 2020;86:54–63. <https://doi.org/10.1016/j.neurobiolaging.2019.09.015>.
5. Kossoff EH, Dorward JL. The modified Atkins diet. Epilepsia 2008;49 Suppl 8:37–41.
6. Krauss RM, Eckel RH, Howard B, Appel LJ, Daniels SR, Deckelbaum RJ, et al. AHA Dietary Guidelines: revision 2000: A statement for healthcare professionals from the Nutrition Committee of the American Heart Association. Stroke 2000;31:2751–66.
7. Gilbert DL, Pyzik PL, Freeman JM. The Ketogenic Diet: Seizure Control Correlates Better With Serum β -Hydroxybutyrate Than With Urine Ketones. J Child Neurol 2000;15:787–90.
8. Marotz C, Amir A, Humphrey G, Gaffney J, Gogul G, Knight R. DNA extraction for streamlined metagenomics of diverse environmental samples. Biotechniques 2017;62:290–3.
9. Marotz C, Cavagnero KJ, Song SJ, McDonald D, Wandro S, Humphrey G, et al. Evaluation of the Effect of Storage Methods on Fecal, Saliva, and Skin Microbiome Composition. mSystems 2021;6. <https://doi.org/10.1128/mSystems.01329-20>.
10. Gonzalez A, Navas-Molina JA, Kosciolk T, McDonald D, Vázquez-Baeza Y, Ackermann G, et al. Qiita: rapid, web-enabled microbiome meta-analysis. Nat Methods 2018;15:796–8.
11. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 2018;34:3094–100.
12. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 2018;34:i884–90. <https://doi.org/10.1093/bioinformatics/bty560>.
13. Zhu Q, Mai U, Pfeiffer W, Janssen S, Asnicar F, Sanders JG, et al. Phylogenomics of 10,575 genomes reveals evolutionary proximity between domains Bacteria and Archaea. Nat Commun 2019;10:5477.
14. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–9.

15. Hillmann B, Al-Ghalith GA, Shields-Cutler RR, Zhu Q, Knight R, Knights D. SHOGUN: a modular, accurate and scalable framework for microbiome quantification. *Bioinformatics* 2020;36:4088–90.
16. Zhu Q, Huang S, Gonzalez A, McGrath I, McDonald D, Haiminen N, et al. Phylogeny-Aware Analysis of Metagenome Community Ecology Based on Matched Reference Genomes while Bypassing Taxonomy. *mSystems* 2022;7:e0016722.
17. Adusumilli R, Mallick P. Data Conversion with ProteoWizard msConvert. *Methods Mol Biol* 2017;1550:339–68.
18. Pluskal T, Castillo S, Villar-Briones A, Oresic M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 2010;11:395.
19. Nothias L-F, Petras D, Schmid R, Dührkop K, Rainer J, Sarvepalli A, et al. Feature-based molecular networking in the GNPS analysis environment. *Nat Methods* 2020;17:905–8.
20. Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat Biotechnol* 2016;34:828–37.
21. Gauglitz JM, Bittremieux W, Williams CL, Weldon KC, Panitchpakdi M, Di Ottavio F, et al. Reference data based insights expand understanding of human metabolomes. *bioRxiv* 2020:2020.07.08.194159. <https://doi.org/10.1101/2020.07.08.194159>.
22. Barabási A-L, Menichetti G, Loscalzo J. The unmapped chemical complexity of our diet. *Nature Food* 2019;1:33–7.
23. Martino C, Morton JT, Marotz CA, Thompson LR, Tripathi A, Knight R, et al. A Novel Sparse Compositional Technique Reveals Microbial Perturbations. *mSystems* 2019;4. <https://doi.org/10.1128/mSystems.00016-19>.
24. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 2019;37:852–7.
25. Martino C, Shenhav L, Marotz CA, Armstrong G, McDonald D, Vázquez-Baeza Y, et al. Context-aware dimensionality reduction deconvolutes gut microbial community dynamics. *Nat Biotechnol* 2021;39:165–8.
26. Vázquez-Baeza Y, Pirrung M, Gonzalez A, Knight R. EMPeror: a tool for visualizing high-throughput microbial community data. *Gigascience* 2013;2:16.
27. Bokulich NA, Dillon MR, Zhang Y, Rideout JR, Bolyen E, Li H, et al. q2-longitudinal: Longitudinal and Paired-Sample Analyses of Microbiome Data. *mSystems* 2018;3. <https://doi.org/10.1128/mSystems.00219-18>.
28. Morton JT, Marotz C, Washburne A, Silverman J, Zaramela LS, Edlund A, et al. Establishing microbial composition measurement standards with reference frames. *Nat Commun* 2019;10:2719.
29. Morton JT, Aksenov AA, Nothias LF, Foulds JR, Quinn RA, Badri MH, et al. Learning representations of microbe–metabolite interactions. *Nat Methods* 2019;16:1306–14.
30. Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res* 2014;42:D459–71.