



OPEN

## Whole-body metabolic modelling reveals microbiome and genomic interactions on reduced urine formate levels in Alzheimer's disease

Filippo Martinelli<sup>1,2</sup>, Almut Heinken<sup>1,2,3</sup>, Ann-Kristin Henning<sup>4</sup>, Maria A. Ulmer<sup>5</sup>, Tim Hensen<sup>1,2</sup>, Antonio González<sup>6</sup>, Matthias Arnold<sup>5,7</sup>, Sanjay Asthana<sup>8</sup>, Kathrin Budde<sup>4</sup>, Corinne D. Engelman<sup>33</sup>, Mehrbod Estaki<sup>6</sup>, Hans-Jürgen Grabe<sup>9</sup>, Margo B. Heston<sup>8</sup>, Sterling Johnson<sup>8</sup>, Gabi Kastenmüller<sup>5</sup>, Cameron Martino<sup>6</sup>, Daniel McDonald<sup>6</sup>, Federico E. Rey<sup>10</sup>, Ingo Kilimann<sup>11,12</sup>, Olive Peters<sup>13,14</sup>, Xiao Wang<sup>14</sup>, Eike Jakob Spruth<sup>13,15</sup>, Anja Schneider<sup>16,17</sup>, Klaus Fliessbach<sup>16,17</sup>, Jens Wiltfang<sup>18,19,20</sup>, Niels Hansen<sup>19</sup>, Wenzel Glanz<sup>21</sup>, Katharina Buerger<sup>22,23</sup>, Daniel Janowitz<sup>23</sup>, Christoph Laske<sup>24,25,26</sup>, Matthias H. Munk<sup>24,26</sup>, Annika Spottke<sup>16,17</sup>, Nina Roy<sup>16</sup>, Matthias Nauck<sup>27,4</sup>, Stefan Teipel<sup>11,12</sup>, Rob Knight<sup>28,29,34,35,6</sup>, Rima F. Kaddurah-Daouk<sup>30</sup>, Barbara B. Bendlin<sup>8</sup>

<sup>1</sup>School of Medicine, University of Galway, Galway, Ireland. <sup>2</sup>The Ryan Institute, University of Galway, Galway, Ireland. <sup>3</sup>Inserm UMRS 1256 NGERE, University of Lorraine, Nancy, France. <sup>4</sup>Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Germany. <sup>5</sup>Institute of Computational Biology, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany. <sup>6</sup>Department of Pediatrics, University of California San Diego, La Jolla, CA, USA. <sup>7</sup>Department of Psychiatry and Behavioural Sciences, Duke University, Durham, NC, USA. <sup>8</sup>Wisconsin Alzheimer's Disease Research Center, School of Medicine and Public Health, University of Wisconsin, Madison, USA. <sup>9</sup>German Center of Neurodegenerative Diseases (DZNE), Rostock/Greifswald, Germany. <sup>10</sup>Department of Bacteriology, University of Wisconsin-Madison, Madison, WI, USA. <sup>11</sup>German Center of Neurodegenerative Diseases (DZNE), Rostock, Germany. <sup>12</sup>Department of Psychosomatic Medicine, University Medicine Rostock, Rostock, Germany. <sup>13</sup>German Center of Neurodegenerative Diseases (DZNE), Berlin, Germany. <sup>14</sup>Department of Psychiatry, Charité-Universitätsmedizin Berlin, Berlin, Germany. <sup>15</sup>Department of Psychiatry and Psychotherapy, Charité, Berlin, Germany. <sup>16</sup>German Center of Neurodegenerative Diseases (DZNE), Bonn, Germany. <sup>17</sup>Department of Neurology, University of Bonn, Bonn, Germany. <sup>18</sup>German Center of Neurodegenerative Diseases (DZNE), Goettingen, Germany. <sup>19</sup>Department of Psychiatry and Psychotherapy, University of Goettingen, Goettingen, Germany. <sup>20</sup>Neurosciences and Signaling Group, Department of Medical Sciences, Institute of Biomedicine (iBiMED), University of Aveiro, Aveiro, Portugal. <sup>21</sup>German Center of Neurodegenerative Diseases (DZNE), Magdeburg, Germany. <sup>22</sup>German Center of Neurodegenerative Diseases (DZNE), Munich, Germany. <sup>23</sup>Institute for Stroke and Dementia Research (ISD), University Hospital, LMU Munich, Munich, Germany. <sup>24</sup>German Center of Neurodegenerative Diseases (DZNE), Tübingen, Germany. <sup>25</sup>Section for Dementia Research, Hertie Institute for Clinical Brain Research, Tübingen, Germany. <sup>26</sup>Section for Dementia Research, Department of Psychiatry and Psychotherapy, University of Tübingen, Tübingen, Germany. <sup>27</sup>DZHK (German Centre for Cardiovascular Research), Partner Site Greifswald, University Medicine, Greifswald, Germany. <sup>28</sup>Department of Computer Science and Engineering, University of California San Diego, La Jolla, CA, USA. <sup>29</sup>Center for Microbiome Innovation, University of California San Diego, La Jolla, CA, USA. <sup>30</sup>Department of Medicine, Duke University Medical Center, Durham, USA. <sup>31</sup>School of Microbiology, University of Galway, Galway, Ireland. <sup>32</sup>APC Microbiome Ireland, Cork, Ireland. <sup>33</sup>Department of Population Health Sciences, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA. <sup>34</sup>Shu Chien-Gene Lay Department of Engineering, University of California San Diego, La Jolla, CA, USA. <sup>35</sup>Halicioğlu Data Science Institute, University of California San Diego, La Jolla, CA, USA. <sup>36</sup>These authors contributed equally: Johannes Hertel and Ines Thiele. ✉email: Johannes.Hertel@med.uni-greifswald.de; ines.thiele@universityofgalway.ie

Johannes Hertel<sup>1,27,36</sup> & Ines Thiele<sup>1,2,31,32,36</sup>

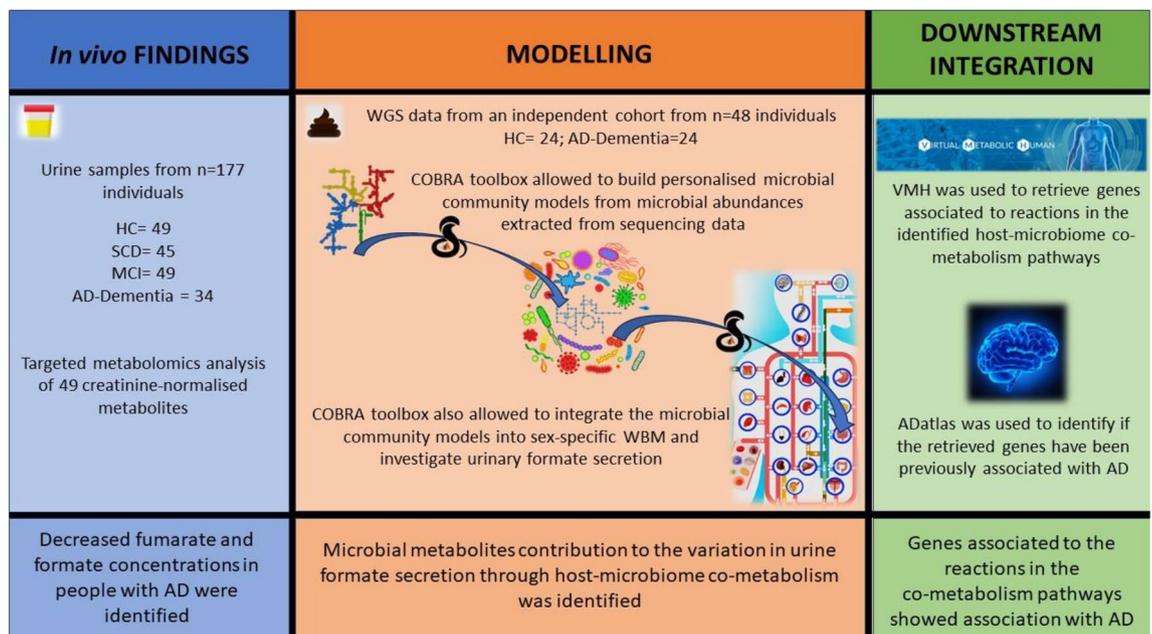
In this study, we aimed to understand the potential role of the gut microbiome in the development of Alzheimer's disease (AD). We took a multi-faceted approach to investigate this relationship. Urine metabolomics were examined in individuals with AD and controls, revealing decreased formate and fumarate concentrations in AD. Additionally, we utilised whole-genome sequencing (WGS) data obtained from a separate group of individuals with AD and controls. This information allowed us to create and investigate host-microbiome personalised whole-body metabolic models. Notably, AD individuals displayed diminished formate microbial secretion in these models. Additionally, we identified specific reactions responsible for the production of formate in the host, and interestingly, these reactions were linked to genes that have correlations with AD. This study suggests formate as a possible early AD marker and highlights genetic and microbiome contributions to its production. The reduced formate secretion and its genetic associations point to a complex connection between gut microbiota and AD. This holistic understanding might pave the way for novel diagnostic and therapeutic avenues in AD management.

**Keywords** Alzheimer's disease, Metabolic modelling, Constraint-based modelling, Microbiome, Formate, Metabolomics, Host-microbiome, Pathways, Co-metabolism

Alzheimer's disease (AD) is the world's leading neurodegenerative disease and the most common cause of dementia<sup>1,2</sup>. Due to ageing populations and the current lack of effective treatment, AD constitutes a major socioeconomic burden in industrialised countries<sup>3</sup>. AD is characterised by progressing brain pathology and declining cognitive function, in addition to metabolic network failures resulting in changes in the metabolome<sup>4</sup>.

The human gut microbiome plays an important role in human health and well-being, performing a wide range of essential functions<sup>5,6</sup>. The gut microbiome affects and is affected by downstream organs, such as the liver, kidney, and brain<sup>7</sup>, and is a major source of neuroprotective (e.g., indoleacrylic acid) and neurotoxic (e.g., quinolinic acid) metabolites<sup>8</sup>. Recently, microbial dysbiosis in AD has been reported<sup>9,10</sup>. Moreover, changes in gut microbial metabolites, such as primary and secondary bile acids (BAs), have been reported to correlate with mild cognitive impairment and AD dementia<sup>11,12</sup>. In particular, gut microbiome contributes up to 50% of formic acid human production<sup>13</sup> and urinary formic acid levels have been recently suggested as non-invasive and cost-effective urinary biomarkers analysis in case of AD<sup>14</sup>.

Creating mechanistic computational models capable of integrating multi-omics data is essential for understanding the connections between shifts in the microbiome, changes in the metabolome, alterations in host metabolism, and the onset of diseases. The constraint-based reconstruction and analysis (COBRA) approach provides a powerful tool to better understand metabolic changes and the biochemical pathways affected by the disease<sup>15</sup>. COBRA relies on manually refined genome-scale metabolic reconstructions that are built bottom-up from the target organisms' genome sequences<sup>16</sup>. These genome-scale reconstructions can be converted into condition-specific metabolic models through the application of constraints, such as dietary information, metagenomics, and metabolomics, thereby enabling the computation of the biologically feasible solution space of these metabolic models through widely used methods, such as flux balance analysis<sup>17</sup>. To meet the scale of human gut microbiomes, a resource of genome-scale reconstructions of 773 human gut microbes, called AGORA<sup>18</sup>, was generated and subsequently expanded in size and scope (AGORA2), now accounting for 7,302 human microbes<sup>19</sup>. Additionally, a large-scale reconstruction resource based on metagenomically assembled genomes has been developed accounting for nearly 250,000 genome-scale metabolic reconstructions<sup>20</sup>. Using these microbial reconstruction resources, comprehensive, personalised microbiome models can be constructed using microbial relative abundances derived from metagenomics data through tools, such as the Microbiome modelling toolbox<sup>21,22</sup>. Using personalised microbiome modelling, it is possible to provide insights and identify



**Figure 1.** Workflow summary *HC* Healthy Controls, *SCD* Subjective Cognitive Decline, *MCI* Mild Cognitive Impairment, *AD* Alzheimer's disease.

changes in the metabolic capabilities of gut microbiomes in disease states, including in inflammatory bowel disease<sup>21,22</sup> and colorectal cancer. Such predicted metabolic capabilities could be validated against faecal metabolome data<sup>23</sup>. Additionally, sex-specific whole-body models (WBMs) of human metabolism have been published<sup>24</sup>, which represent human metabolism in an organ-resolved and anatomically accurate manner. These WBMs can be contextualised with physiological parameters, dietary input, and personalised microbiome models<sup>24</sup>. For the first time, these models can also be used to investigate the interaction between the host and microbiota metabolisms<sup>24,25</sup> and host-virus co-metabolism<sup>26</sup>. These simulation-derived novel hypotheses can be subsequently validated with targeted in vitro or in vivo studies.

## Results

In this study, we aimed to characterise metabolic changes associated with cognitive decline and AD as well as to elucidate the role of the human gut microbiota in these changes. We first generated 1D-NMR urine metabolome data from individuals on the clinical AD spectrum and from cognitively healthy age-matched individuals from the Germany-based DELCODE study cohort. We identified a statistically significant decrease in formate and fumarate concentrations in people with AD dementia. To dissect potential microbial contributions, we then used WGS data generated from an independent cohort consisting of 24 participants with AD dementia and 24 cognitively unimpaired age-sex-BMI-matched controls to construct personalised host-microbiome models for each individual. We found that the formate microbial secretion was lowered in AD host-microbiome models compared to healthy control host-microbiome models. Furthermore, we found that amino acids and sugar host-microbiome co-metabolism were involved in the predicted urine formate production. The implications of these two host pathways in AD were further supported by known associations of the underlying genes with AD pathology. Taken together, our analyses point towards alterations in the microbiome's formate secretion capacity between AD cases and controls and host-microbiome interactions on its production (Fig. 1). These findings have potential clinical relevance for understanding the interplay between the host and the microbiome in AD.

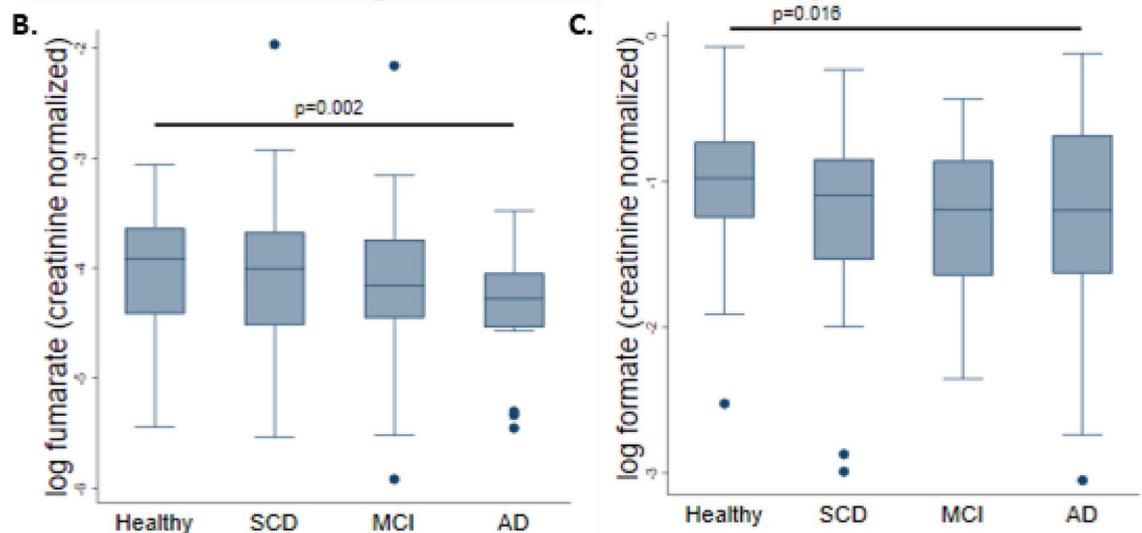
### Metabolomic analysis reveals decreased fumarate and formate concentrations in the urine of people with AD

To characterise altered metabolism in AD and cognitive decline, we utilised 1D-NMR urine metabolome data generated at the baseline visit from n = 177 DELCODE participants<sup>27</sup> (Fig. 2A). The DELCODE cohort has four principal study groups, from which a subsample was characterised using urine metabolomics: healthy controls (n = 49), individuals with subjective cognitive decline (n = 45), mild cognitive impairment (n = 49), and AD-Dementia (n = 34). A targeted metabolomics analysis resulted in the quantification of 49 creatinine-normalised urine metabolites (Table S01). We screened all quantified metabolites with at least 50% non-zero measurements in association with the study group variable. The tricarboxylic acid (TCA) cycle intermediate fumarate was decreased in the urine among people with AD dementia (FDR < 0.05, Fig. 2B, Table S02). Furthermore, formate was also decreased in the AD dementia group, although the result was not statistically significant after correction for multiple testing (Fig. 2C, Table S02). Notably, formate and fumarate were lowered among individuals with subjective cognitive decline (SCD) as well as those with mild cognitive impairment (MCI), suggesting that altered formate and fumarate could be indicative of an early process in the aetiology of AD.

## A. Descriptive statistics for the analysed DELCODE samples

	Healthy (n=49)	SCD (n=45)	MCI (n=49)	AD (n=34)	p-value
Age, mean (SD)	68.02(5.32)	70.84(6.12)	72.39(5.27)	71.71(6.78)	0.002*
BMI, mean (SD)	26.06(3.59)	26.47(3.98)	25.73(3.33)	24.21(3.17)	0.038*
Female, %	57.14%	55.56%	63.27%	41.18%	0.119 <sup>#</sup>
Smoking, %	10.2%	11.11%	4.08%	8.82%	0.594 <sup>#</sup>

## B.-C. Results of targeted metabolome analyses



**Figure 2.** 1D-NMR metabolomic data on urine samples from the DELCODE study cohort support altered metabolism in AD, including formate metabolism. (A) Descriptive statistics for the analysed DELCODE samples. \**p*-value derived from one-factorial ANOVA, <sup>#</sup>*p*-value derived from Fisher's exact test. SCD subjective cognitive decline, MCI mild cognitive impairment, AD Alzheimer's disease, BMI body mass index. (B, C) Box plots for formate and fumarate (both creatinine normalised) over the four study groups. *p*-Values were derived from multivariable regressions adjusting for age, sex, and BMI using heteroscedastic robust standard errors. SCD subjective cognitive decline. MCI mild cognitive impairment. AD Alzheimer's disease.

### Microbiome models showed no difference in metabolic content

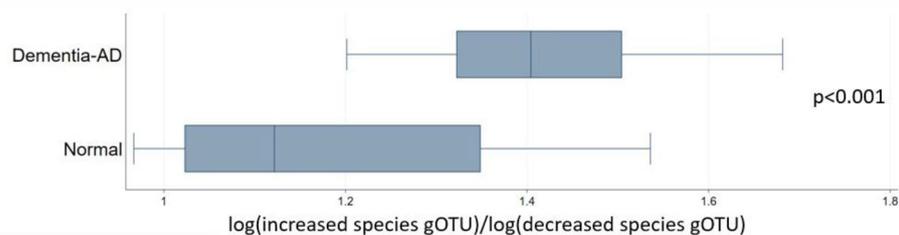
The human gut microbiome affects host metabolism and contributes up to 50% to human formate production<sup>13</sup>. To investigate possible microbial influences on the altered fumarate and formate levels in urine, we used personalised constrained-based models of human and microbial metabolism. As no microbiome data for the DELCODE study participants were available, we obtained stool samples from 24 subjects with dementia due to AD and 24 age- and sex-matched control participants from the Wisconsin Registry for Alzheimer's Prevention study (WRAP) and the Wisconsin Alzheimer's Disease Research Center cohort (WADRC)<sup>28</sup>. As expected, there was a higher carriage of the *APOE4* allele among people with AD dementia (Fig. 3A), the strongest genetic risk factor for late-onset AD<sup>29,30</sup>. For these 48 individuals, metagenomic sequence data were generated from the stool samples and analysed using the default Woltka<sup>31</sup> toolkit against Web of Life (Release 1)<sup>32</sup> in Qiita<sup>33</sup>. The median amount of reads the samples contained was 2,561,443 (interquartile range (IQR) = 2,344,948) reads. (Fig. 3A). No statistically significant differences in the number of reads could be observed between cases and controls.

To gain insight into microbiome metabolism, we used the microbial species' relative abundances to generate personalised microbiome metabolic models and we investigated their properties. Therefore, we discarded all reads that had no species taxonomic information, compared the list of identified species with the ones accounted for by the AGORA2 metabolic reconstruction resource<sup>19</sup>, and discarded all reads belonging to species that could not be mapped onto AGORA2. We generated pan-species microbial metabolic reconstructions consisting of the union of all metabolites and reactions present in any corresponding AGORA2 strains of the same species ("Methods"). These pan-species microbial reconstructions were then combined into one microbiome reconstruction for each sample. This AGORA2 mapping resulted in 815 different species present in at least one of the 48 samples with an average of 476 (SD = 74) microbes in each sample. After the cut-off of low abundant species ("Methods"), the sample-specific microbiome models were composed of an average of 259 (SD = 29) pan-species models, accounting for an average of 73.60% (SD = 11.92%) of the total reads in the sequencing data (Fig. 3A). The resulting 48

## A. Samples and models characteristics of the Wisconsin cohort study

	Participants (n=48)	Dementia-AD (n=24)	Healthy (n=24)	p-value
Age, mean (SD)	70.318 (7.521)	71.409 (7.398)	69.228 (7.641)	0.320 <sup>a</sup>
APOE4, %	50.00%	79.17%	20.83%	<0.001 <sup>b</sup>
Total reads, mean (SD)	5,849,867 (9,661,968)	6,797,641 (10,318,332)	4,902,094 (9,078,737)	0.503 <sup>a</sup>
AGORA2 mapped reads, mean (SD)	4,371,723 (7,246,048)	5,136,296 (7,430,138)	3,607,150 (7,132,063)	0.471 <sup>a</sup>
Metagenomic data species, mean (SD)	1,131 (336)	1,132 (386)	1131 (286)	0.995 <sup>a</sup>
AGORA2 mapped species, mean (SD)	476 (74)	478 (81)	475 (68)	0.905 <sup>a</sup>
Species in the model after cut-off, (SD)	259 (29)	251 (24)	267 (31)	0.058 <sup>a</sup>
Reactions in microbiome models, mean (SD)	373,118 (40,819)	362,678 (34,281)	383,559 (44,745)	0.076 <sup>a</sup>
Metabolites in microbiome models, mean (SD)	337,341 (36,718)	327,857 (30,863)	346,824 (40,185)	0.073 <sup>a</sup>
Unique reactions in microbiome models, mean (SD)	7,559 (169)	7,527 (159)	7,591 (176)	0.193 <sup>a</sup>

## B. Log-ratio analysis on altered microbes



**Figure 3.** Descriptive statistics for the analysed samples and their corresponding Qiita-derived microbiome models, and diversity analysis on microbial relative abundance. **(A)** Samples and model characteristics of the Wisconsin cohort studies. AD Alzheimer's disease, SD Standard deviation, <sup>a</sup>p-value from Welch t-tests, <sup>b</sup>p-value from Fisher's exact test. Full results can be found in Table S03. **(B)** Boxplot of gOTU log-ratio analysis on microbes, whose models' relative abundances were found to be altered between healthy and dementia-AD participants, p-value from Welch t-tests. The full results of the relative abundance analysis can be found in Table S04.

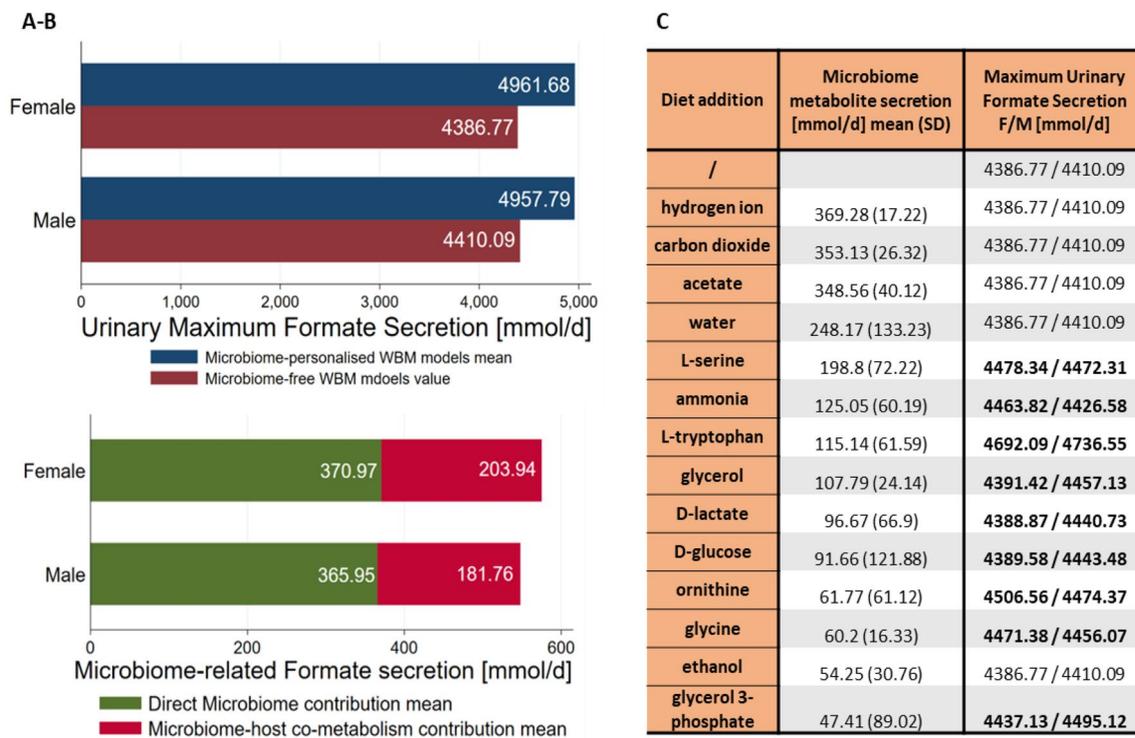
sample-specific microbiome models consisted, on average, of 373,118 (SD = 40,819) non-unique metabolites and 337,340 (SD = 36,718) non-unique reactions (Fig. 3A). The reaction and metabolite redundancy in the microbiome models originated from the fact that each pan-microbial reconstruction was preserved within the microbiome reconstructions. When only considering the set of unique reactions, the AD microbiome models contained on average 7,558 (SD = 169) unique reactions. No differences in the metabolic content were detected between AD and control microbiome models (Fig. 3A).

## Microbes' abundance shifts in AD-dementia

We investigated altered microbes' relative abundances in the microbiome reconstructions and found that the relative abundances of *Turicibacter sanguinis*, *Hungatella hathewayi*, *Turicibacter sp h121*, *Haemophilus parainfluenzae*, *Ruminococcus champanellensis*, *Dialister succinatiphilus*, and *Coprobacillus cateniformis* were decreased and the relative abundances of *Akkermansia sp kle1797*, *Akkermansia sp kle1798*, *Parabacteroides distasonis*, *Bacteroides finegoldii*, *Collinsella tanakaei*, *Collinsella stercoris*, and *Bacteroides thetaiotaomicron* were increased in the AD-Dementia (participants with dementia due to AD) samples compared to the healthy controls (Table S04). Due to the compositional nature of the relative abundance data, we performed a log ratio analysis of the increased species against the decreased ones using their gOTU counts. This analysis confirmed the alteration found in the relative abundances (Fig. 3B).

## No alteration in formate urine secretion fluxes was predicted between AD dementia participants and healthy controls

Formate is a fermentation production of anaerobic bacteria in the gut and the gut microbiome metabolism can account for up to 50% of the overall formate production in the host<sup>13</sup>. To evaluate the possible involvement of the microbiome on formate urine excretion, we used the sex-specific, organ-resolved whole-body models of human metabolism (WBM)<sup>24</sup> and added the personalised microbiome models to the large-intestinal lumen of either the female or male WBM. Formate, which we found altered in the urine of people with AD in the DELCODE study, was part of the microbiome models and the WBMs. This metabolic overlap enabled our investigation of its direct effect on host-microbiome co-metabolism. We computed the maximum urine secretion flux and the maximum microbiome production flux for formate for all the microbiome-associated WBMs and the two sex-specific germ-free WBMs. The simulation setup of the WBMs was identical except for sex-specific host metabolism and the microbiome models. Hence, we included sex as a covariate in the subsequent statistical analyses. No statistically significant differences could be observed between healthy and AD individuals (Table S05).



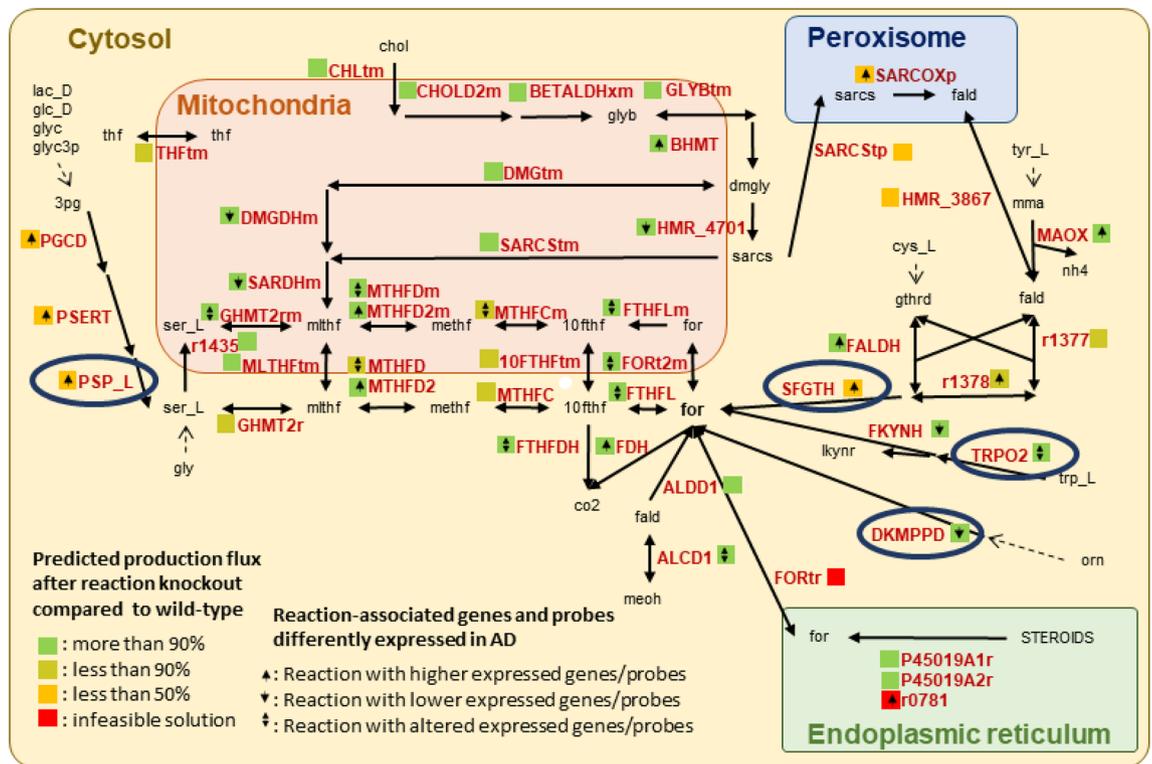
**Figure 4.** Evaluation of host-microbiome involvement in urine secretion of microbe-derived metabolites. (A, B) Breakdown of maximum sex-specific urine formate secretion highlighting the presence of microbiome-host co-metabolism. (C) Average metabolite microbiome secretion when the host-microbiome WBMs were interrogated for the maximum urine secretion, and urine secretions when the specific metabolites were unconstrained in the diet for the germ-free models.

### Microbial metabolites contribute to the variation in urine formate secretion through host-microbiome co-metabolism

To evaluate the presence and contribution of host-microbial co-metabolism in formate urinary production, we investigated the predicted differences between the formate urine fluxes of the microbiome-personalised WBMs and non-microbiome-personalised WBMs, i.e., germ-free WBMs. The difference was higher than the amount of formate secreted by the microbiomes meaning that other microbially derived metabolites must be involved in the host-microbiome co-metabolism of formate (Fig. 4A,B). To determine the main microbial metabolites that contributed to urinary formate production, we investigated the solution vector when solving the microbiome-personalised models for maximum formate urine secretion and extracted all the microbial secretion fluxes that had average flux values greater than 30 mmol/person/day (TableS06). Excluding formate, 14 metabolites were found to be highly secreted by the microbiomes. To test whether these metabolites could indeed increase the maximal urine formate secretion, we added each of the 14 metabolites individually to the diet of germ-free WBMs and solved for the maximum formate urine secretion. Nine of the 14 metabolites resulted in higher flux values for urine formate secretion compared to the average European diet without their addition and thus could be used by the host metabolism to synthesise formate (Fig. 4C). Interestingly, one of the metabolites, the ammonium ion, is not a carbon source. Nonetheless, its addition to the diet increased the formate urine secretion suggesting its involvement in formate-producing pathways. These results demonstrate that the host models metabolise microbial secretion products underlining the importance of microbial-derived metabolites as precursors of urinary formate.

### Host formate originates from the metabolism of dietary and microbial metabolites

Subsequently, we aimed at elucidating the host-microbiome metabolism pathways involved in formate production. The host can produce formate through various pathways from dietary inputs<sup>13,34</sup>. We considered 47 reactions that have been suggested to be implicated in host formate metabolism and three transport reactions into or from the peroxisome and the endoplasmic reticulum to complete the pathways (Fig. 5, Table S07)<sup>13</sup>. To quantify the contribution of host metabolic reactions involved in the urine formate secretion, we deleted metabolic reactions involved in formate production in the germ-free WBM models. Overall, 17 of the 47 (36%) reaction deletions reduced the urine formate secretion flux by more than 10% (Fig. 5, Table S08). We then identified a minimal set of four reactions (VMH ID: PSP\_L, Phosphoserine phosphatase; VMH ID: SFGTH, S-Formylglutathione hydrolase; VMH ID: TRPO2, L-Tryptophan: Oxygen 2,3 Oxidoreductase; and VMH ID: DKMPPD, 2,3-Diketo-5-Methylthio-1-Phosphopentane Degradation) (Fig. 5), which, when deleted together, largely reduced the urine formate secretion by the germ-free WBM models (92% reduction in male, 86% in female). These four reactions represent endpoints of multiple pathways leading eventually to formate production, either directly or indirectly



**Figure 5.** Cellular metabolism involved in the production of formate including reactions found responsible for host-microbiota co-metabolism. Dotted lines represent diet constituent involvement in the overall formate production, circled reactions when deleted together largely reduced the formate urinary production. Metabolites names (abbreviations are given in VMH IDs, [www.vmh.life](http://www.vmh.life),<sup>69</sup>): *10fth* 10-formyl-THF; *3pg* 3-phosphoglyceric acid; *chol* choline; *cys\_L* L-cysteine; *dmgly* dimethylglycine; *fald* formaldehyde; *for* formate; *glc\_D* D-glucose; *gly* glycine; *glyb* betaine; *glyc* glycerol; *glyc3p* glycerol-3-phosphate; *gthrd* GSH; *his\_L* L-histidine; *lac\_D* D-lactate; *lkynr* L-kynurenine; *meoh* methanol; *methf* 5,10-methenyl-THF; *mlthf* 5,10-methylene-THF; *mma* methylamine; *orn* ornithine; *sarcs* sarcosine; *ser\_L* L-serine; *thf* tetrahydrofolate; *trp\_L* L-tryptophane; *tyr\_L* L-tyrosine. Reactions' names can be found in Table S07. Complete KO analysis can be found in Table S08. Differential gene expression (DEG) and differential protein abundance (DEP) analysis of association to AD-related phenotypes can be found in Table S09.

via the formation of formaldehyde, which is then converted into S-formylglutathione by formaldehyde dehydrogenase (VMH ID: FALDH) and formate by the S-formylglutathione hydrolase (VMH ID: SFGTH) (Fig. 5). The four reactions were present in numerous organs in the WBM models, with all being present in the brain, pancreas, lung, and adrenal gland. Different subsets of three were also present in eight further organs (colon, stomach, liver, heart, adipocytes, gall, kidney, and thyroid gland). Notably, the four reactions were also part of the metabolic pathways of the ten microbial metabolites that we previously identified (Fig. 4C). Taken together, microbial metabolites contributed to urine formate secretion flux via host-microbiome co-metabolism involving a network of host reactions that could be mostly deactivated by the deletion of four reactions, which are broadly distributed across numerous organs.

### Reactions involved in the urine formate secretion are associated with AD

We were interested in whether any of the identified reactions and their corresponding genes have been previously associated with AD. Using the ADAtlas (<https://adatlas.org/>)<sup>35</sup> a network-based integrative resource for AD, we found that 28 of the 47 (59.5%) reactions had genes associated with AD/cognitive decline phenotypes (Fig. 5, Table S09). Nine of these 28 reactions (32%) were also predicted to alter the maximal urine formate secretion flux by more than 10% (Fig. 5). Six of these nine reactions (i.e., VMH IDs: SARCOXp, r1378, SFGTH, PGCD, PSERT, and PSP\_L, full names in Table S07) were involved in the formate production through formaldehyde and 3-phosphoglycerate catabolism and positively associated with AD. Additionally, two reactions (VMH ID: FKYNH, N-formyl-L-kynurenine amidohydrolase; and VMH ID: TRPO2, L-tryptophan: oxygen 2,3-oxidoreductase) were involved in formate production through tryptophan catabolism and negatively associated with AD. Notably, reactions associated with the folate metabolism (MTHFDm, MTHFDm2, MTHFCm, MTHFD, MTHFD2) were also associated with genes differently expressed in AD, in particular with low-expression of gene *MTHFD1* and high-expression of gene *MTHFD2* (Fig. 5, Table S09). These results demonstrate that formate metabolism is altered in people with AD due to genetic variations and these genetic changes could lead to changes in formate urine secretion in addition to microbial formate production.

## Discussion

A challenge in understanding the gut microbiome's role in health and neurological disease is linking changes in the microbiome itself and its derived metabolome to changes in host metabolism and cognitive symptoms. Here, we present a personalised metabolic modelling approach for the interrogation of the host-microbiome co-metabolism in AD. Through this framework, we used whole-body models of human metabolism to generate insights about the consequences of altered formate microbial secretion and its association with urine output through host-microbiome co-metabolism. These *in silico* investigations were motivated by findings of alterations in formate urine levels in AD from an independent cohort. Taken together, our analyses point towards altered gut microbiome secretion capacities and host-microbiome interactions on formate production, results that could help understand the crosstalk between host and microbiome metabolisms in AD.

The use of urine metabolomics, as a non-invasively collectable biofluid, in the context of AD and the identification of early markers of AD could be crucial to developing successful therapies<sup>36,37</sup>. Notably, urinary formate has been already suggested as a new potential biomarker for Alzheimer's disease by an independent study<sup>14</sup>. As a final breakdown product of human and microbial metabolism, formate is typically found in human urine<sup>38</sup>. Levels are affected by both environmental and dietary exposures, and our results suggest that the urinary formate secretion could be a direct effect of altered host and microbial formate co-metabolism<sup>13</sup>. The interrogation of microbiome-personalised sex-specific WBM models, through computation of maximum formate urine secretions (Fig. 5), highlighted that microbial formate secretion capacity was lower in AD microbiomes. Our results suggested the involvement of host-microbiota co-metabolism in the overall formate urine secretion catabolising microbially derived metabolites, such as glucose, L-serine, L-glycine, L-tryptophan, L-cysteine, L-tyrosine, and ornithine. The tyrosine pathway has been repeatedly implicated in AD<sup>39,40</sup> and tryptophan metabolites have been shown to regulate the cerebral activity of neprilysin, a metalloproteinase that controls the degradation and clearance of A $\beta$  peptides in the brain<sup>41</sup>.

Our results further underline altered metabolism as a hallmark of AD<sup>42,43</sup>. In particular, the WBMs revealed the role of amino acid degradation in formate production, highlighting microbial-derived tryptophan degradation as one of the primary microbial sources of formate. Tryptophan is not only an important precursor of neurotransmitters and neuroactive metabolites, such as serotonin and kynurenine<sup>44</sup>, but it also plays a role in immunoregulation<sup>4</sup>. Moreover, tryptophan depletion increases cognitive deficits among people with AD<sup>45</sup> and the bioavailability of metabolites in the serotonin and kynurenine pathways are altered in both the urine and serum of AD patients<sup>46</sup>. The microbiome modelling implicated that microbial tryptophan production may also be reduced in AD, concurring with earlier work indicating that the microbiome contributes to human tryptophan pools<sup>24</sup>. Thus, in conjunction with the WBMs, our finding of decreased formate among individuals with AD and MCI suggests alterations in tryptophan degradation in AD. These results are also in line with a recently formulated hypothesis of AD being a tryptophan metabolism-correlated disease<sup>47</sup>.

Formate is also involved in different pathways and is a precursor of purine synthesis<sup>48</sup>. Our study also highlighted the association of formate with folate metabolism, a pathway that has been found associated with AD and DNA methylation<sup>49</sup>. The importance of host-microbiome formate co-metabolism is further highlighted by our examination of genes associated with AD, where we found that most of the reactions involved in formate metabolism (Fig. 5) belong to genes, whose expression was altered in AD participants compared to healthy controls. Five reactions associated with the folate metabolism (MTHFDm, MTHFDm2, MTHFCm, MTHFD, MTHFD2) were associated with genes differently expressed in AD, three reactions (VMH IDs: PGCD, PSERT, and PSP\_L), which are involved in the catabolism of 3-phosphoglycerate through an alternative pathway from glycolysis with the production of L-serine, a possible precursor of formate, were associated with genes overexpressed in AD; this result could corroborate the reported reduction in glycolysis intermediate concentrations in AD participants<sup>50</sup>. Overall, these results suggest that formate metabolism is altered in individuals with AD also due to genetic variations, which could lead to changes in formate urine secretion in addition to microbial formate production. Notably, this inference would have not been possible without the WBM modelling, which clarified the role of the host formate metabolism. Without this additional *in silico* analysis, one could have falsely concluded that changes in microbial formate production, due to differences in microbiome composition, would be responsible for the reduced urine formate secretion in AD patients as measured in the metabolomic data. Since our WBM models were not further personalised using an individual's genomic, metabolomic, or transcriptomic data, the aforementioned host genetic factors should be considered in future *in silico* studies, potentially increasing the validity of the *in silico* results regarding host urinary formate secretion.

While the COBRA modelling approach is a very valuable approach for investigating host-microbiome co-metabolism involvement in AD, certain limitations should be noted. For instance, we used a relatively small cohort of subjects and controls. Hence, our results need to be validated in larger independent cohorts with possibly urine metabolomics and WGS data from the same individuals to reduce variability. It has to be noted that discarded microbes not accounted for by the AGORA2 resource could lead to loss of metabolic capacities in the correspondent microbiome models. Additionally, the models were built on microbial relative abundance data generated from the reads count data, and results were subject to the intrinsic compositional structure of the models. Genome-scale metabolic reconstructions are also continuously updated as new experimental data and biochemical knowledge become available<sup>51–53</sup>. The incompleteness of the metabolic reconstructions is particularly true for gut microbes, for which only limited data are available. Computational reconstruction tools, such as DEMETER<sup>54</sup>, which has been used for the construction of the AGORA2 microbial reconstructions, permit the inclusion of experimental data, e.g., from BacDive<sup>55</sup>, during the reconstruction process. Similarly, refined genome annotations that correct missing and mis-annotations should be performed to minimise the errors in the reconstruction, and thus increase the fidelity of the predictions. Such reannotation has been done for most of the microbes in AGORA2. Moreover, an inherent limitation of the COBRA approach is that it assumes the

biological system to be in a steady-state condition, thereby ignoring the dynamic nature of microbial communities. Notably, the predicted secretion capacities obtained through microbiome modelling are not confounded by different factors, such as age and sex, improving the identification of secretion-microbial correlations, since they are derived from deterministic modelling rather than inferred from statistical dependence patterns of observational data. Additionally, while this study highlighted the role of gut microbes and host metabolism and genetics, differences between lifestyle factors (e.g., diet and exercise) and medications are also likely to contribute to changes in host-microbiome co-metabolism, urine metabolome, and AD pathology.

In conclusion, in this study, we combined omics data with COBRA modelling on the level of the microbiome and the whole human supra-organism and highlighted the role of microbiome-host interplay on formate-producing pathways. In particular, the microbiome's role in linking aminoacidic and glucose metabolism with formate, a possible early marker for AD, could be of clinical importance, potentially contributing to the AD phenotype. The underlying mechanism suggested by our model, that both gut microbes and host genetics contribute to an altered formate metabolism in AD, needs to be assessed with more targeted validation studies. Our study delivers proof of the concept of personalised whole-body modelling in the context of a complex human disease. As such, the paradigm has demonstrated promise in uncovering host-microbiome co-metabolism involving biomarkers found in metabolomic studies validating or suggesting pathology hypotheses.

## Methods

This study did not generate new unique reagents.

### Statistical analyses

For descriptive statistics, metric variables were expressed in means and standard deviations, while categorical variables were described by proportions. All p-values were reported as two-tailed. The statistical analyses were performed with STATA 16.1/MP (STATA Inc., College Station, Texas, USA).

### DELCODE cohort

#### *Study sample*

We used an interim data-freeze from the DELCODE study conducted by the Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE)<sup>27</sup>. Note that the DELCODE study did not include individuals with a current major depressive episode, major psychiatric disorders, neurological diseases other than AD, or unstable medical conditions<sup>27</sup>. For our study, we also excluded individuals who had no urine samples available. The resulting sub-cohort consisted of 49 healthy controls, 45 cases with subjective cognitive decline (n = 45), 49 cases with mild cognitive impairment, and 34 cases with AD dementia (Fig. 2A). We defined subjective cognitive decline (SCD) as a constant self-perceived cognitive decline without observation of any objective cognitive impairment as measured by the CERAD test battery, persisting at least for six months and being unrelated to an episodic event<sup>56</sup>. The National Institute on Aging-Alzheimer's Association (NIA-AA) workgroup guidelines<sup>57,58</sup> were used to define the core clinical criteria for MCI and dementia AD-related. The control subjects showed no objective cognitive impairment in cognitive tests, had no history of neurological or psychiatric disease, and did not report a self-perceived cognitive decline. Informed consent was provided by all participants or their representatives. The study protocol was approved by the local institutional review boards and ethics committees of the participating centres and the study was conducted in accord with the Helsinki Declaration of 1975.

#### *<sup>1</sup>H-NMR measurements*

The urine samples were handled and prepared as in<sup>59</sup>. The urine samples from 185 individuals were measured at the University Medicine Greifswald on a Bruker AVANCE-II 600 NMR spectrometer operated by TOPSPIN 3.2 software (both Bruker Biospin, Rheinstetten, Germany). The spectrometer was equipped with a 5-mm z-gradient probe and an automated tuning and matching (ATMA) unit (both Bruker Biospin, Rheinstetten, Germany). Specimens were automatically delivered to the spectrometer via SampleJet (Bruker Biospin, Rheinstetten, Germany) into standard 5 mm NMR tubes. The acquisition temperature was set to 300°K. A standard one-dimensional <sup>1</sup>H-NMR pulse sequence with suppression of the water peak (NOESYPREAST) was used<sup>60</sup>. Urine samples of four individuals could not be processed at all due to too little biomaterial available for <sup>1</sup>H-NMR measurements. After processing the raw spectra, 50 metabolites were quantified using the Bruker Suite B.I.Quant-UR b™ for targeted analyses. To account for dilution, two different normalisation approaches were performed: (1) creatinine normalisation and (2) PQN normalisation. Different normalisation approaches have been shown to have different advantages as well as disadvantages such that it is recommended to use multiple approaches and to compare results across normalisation techniques<sup>61</sup>. For quality check purposes, creatinine was measured via an enzymatic standard kit to compare <sup>1</sup>H-NMR-derived creatinine measurements with standard creatinine measurements. Four further observations were excluded for strong differences between enzymatic and <sup>1</sup>H-NMR creatinine measurements. The remaining observations showed a correlation > 0.95 between the two types of creatinine measurements. In the end, n = 177 spectra were included in statistical analyses.

#### *Descriptive sample statistics of targeted metabolome data*

Metadata and urine samples for 177 individuals belonging to four study groups were available for analysis (Fig. 2A). As described above, eight individuals were excluded from analyses, thus the final sample, for which urine metabolome data was available, comprised 177 individuals. Figure 2A lists the descriptive statistics for basic sample characteristics. Importantly, the four study groups (healthy, subjective memory impairment, MCI, AD) were not balanced for age and body mass index (BMI, Fig. 2A). Therefore, all statistical metabolome analyses accounted for age and BMI differences by including age and BMI as covariates in the statistical modelling.

### *Statistical analyses of the NMR metabolome data*

Descriptive statistics were derived for 49 metabolites after creatinine normalisation. Of those 49 metabolites, only 16 metabolites had non-zero urinary concentration measurements in more than 50% of the samples. Importantly, a zero measurement in the context of NMR measurement does not necessarily mean that the metabolite was below the limit of detection. The signal of the correspondent metabolite may also have been clouded by other metabolites or may have been shifted due to pH differences in ion concentrations, hence, we treated zero measurements as missing values. We screened all metabolites fulfilling the 50% criteria via multivariable linear regressions including age, sex, and BMI as covariates. Log-transformed, creatinine-normalised metabolite concentrations were used as the dependent variables, and the study group variables (categorical) were the predictors of interest. We performed the global test using the Wald-test method to investigate 1. whether the study group variable contributed overall to the statistical model and 2. whether there were specific differences between the AD group and the healthy controls. For correction of multiple testing, the false discovery rate was utilised.

## **WISCONSIN cohort**

### *Faecal samples*

Our study utilised data from the Wisconsin Alzheimer's Disease Research Center (WADRC) and the Wisconsin Registry for Alzheimer's Prevention (WRAP) studies. Participants with AD-Dementia were recruited from the WADRC, while cognitively unimpaired participants were recruited from the WADRC and the WRAP studies. Data on age, sex, BMI, *APOE* genotype, diagnosis, and cognitive test results were collected<sup>62,63</sup>. Our study used a subset of the cohort comprising 48 individuals (24 AD-Dementia cases and 24 sex and age-matched healthy controls). Covariates were compared against the clinical diagnosis using Fisher's exact test for *APOE4* and sex, and Welch's t-test for age and BMI. The study protocol was approved by The University of Wisconsin Health Sciences Institutional Review Board.

### *Sequencing and processing of metagenomic data*

The faecal samples were sequenced and then processed through Qiita<sup>33</sup> using the default workflow for metagenomics data. In short, the raw files were loaded in multiple preparations to represent the multiple runs and processed with default parameters; then the raw sequencing data were demultiplexed and trimmed at 150 bases. Adapter removal was carried out using fastp<sup>64</sup> and human reads were filtered using minimap2<sup>65</sup>. Genomic OTUs (gOTUs) were generated using the Woltka Toolkit<sup>31</sup> by aligning reads to the Web of Life (Release 1) reference genome database<sup>32</sup> using bowtie2<sup>66</sup>.

### *Construction of personalised microbiome metabolic models*

First, we assigned taxonomy to gOTUs using the Web of Life database, and we summed together the gOTUs with matching taxonomical information. Reads with no specific species information or associated to species not accounted by our database of genome-scale metabolic microbial reconstructions (AGORA2)<sup>19</sup> were discarded. The remaining reads were normalised such that the sum of all reads was one, thereby obtaining the relative abundances of each microbe in the sample. The obtained relative abundances were then mapped to 1,742 AGORA2 pan-species reconstructions, the union of reactions and metabolites of each strain-specific reconstruction of one species. Then, sample-specific microbiome models were derived using the mgPipe module of the Microbiome Modelling Toolbox<sup>21,22</sup>. The process involves integrating various microbial reconstructions by introducing additional reactions into each microbial reconstruction to allow the transfer of metabolites between the extracellular space of each microbe and a shared lumen compartment<sup>67</sup>. This lumen compartment is interconnected with both dietary and faecal compartments, allowing for the absorption and excretion of metabolites to and from the surrounding environment. Within each microbial community model, adjustments are made to the community biomass reaction based on the provided relative abundance data. Further constraints were then applied to couple each microbial reaction to its corresponding biomass reaction. These 'coupling constraints' ensure that flux through a pan-genome model within the microbiome model was only non-zero if the corresponding biomass reaction carried a non-zero flux. The coupling factor was arbitrarily chosen to be 400<sup>67,68</sup>. We tested for differences in measures of metabolic content in the microbiome models between cases and controls using linear regressions correcting for the presence of the *APOE4* allele.

### *Species abundances evaluation*

To investigate differences in the AGORA2-covered microbial abundances between cases and controls, we fitted fractional regressions using the relative microbial abundances as the response variable and the health status as an independent one correcting for categorical *APOE* genotype, while we excluded all the microbes present in less than 50% of the models to avoid statistical artefacts. Alteration in relative species abundances was checked using log-ratio analysis, after calculating the sum of the gOTUs for the increased species and decreased species, the ratio was calculated and a t-test was run between groups.

### *Diet constraints*

An average European Diet<sup>69</sup>, supplemented with bile acids (cholic acid and chenodeoxycholic acid), was applied to further constrain each microbiome model and to convert them into condition-specific models (Table S11). The diet constraints were defined to be in mmol/person/day. We integrated the uptake fluxes values defined by the diet with all the microbiome models using the Microbiome Modelling Toolbox<sup>21</sup> implemented in the COBRA toolbox<sup>70</sup>, and we ensured that all the pan-species models could grow under the defined diet.

### Microbiome model simulations

Each microbiome model under the given condition-specific constraints was interrogated using flux variability analysis (FVA)<sup>71</sup> to obtain the corresponding maximum net secretion capacity for all the model's secretion reactions, FVA solves maximisation problems with the interested reactions secretion fluxes as objective subject to different constraints to specify the feasible region: predetermined upper and lower bound of reactions fluxes stated in the model, diet fluxes, and mass balance in form of  $(S \cdot v = dv/dt = 0)$ , whereas is the stoichiometric matrix with metabolites as rows and reactions as columns and  $v$  is the flux vector). For each metabolite secretion capacity, zero values meant that the model could not secrete the corresponding metabolite, while a positive value corresponded to the microbiome model's capacity of secreting the metabolite, under the given simulation constraints. All simulations were performed in MATLAB (Mathworks, Inc.) version R2018b using IBM CPLEX (IBM, Inc.) as a linear programming solver and the COBRA Toolbox v3<sup>70</sup> and the Microbiome Modelling Toolbox<sup>22</sup>.

### Interrogation of whole-body metabolic (WBM) personalised models

We added the individual microbiome models to the large intestinal lumen of either female or male organ-resolved, whole-body models of human metabolism (WBMs), as appropriate, using the PSCM toolbox<sup>24</sup>. We computed the maximum urine secretion fluxes of metabolites, which were significant in the previous analyses and could be produced by the microbiome-associated WBM under an average European diet<sup>69</sup>. The male and female WBMs, which correspond to a reference man and woman<sup>24</sup>, respectively, were not further personalised as the sub-cohort was age and BMI matched between healthy and AD individuals. The maximum urine secretion rate for each microbiome-associated WBM was calculated by choosing the corresponding urine metabolite exchange reaction (e.g., 'EX\_for[u]' for formate) as an objective function and maximising for this reaction, using the PSCM toolbox<sup>24</sup>. Maximum urine secretion rates for the investigated metabolites were also calculated for non-personalised, germ-free, male and female WBMs. The maximum microbial secretion fluxes were calculated by choosing the corresponding microbial metabolite exchange reaction between the microbial lumen [luM] and large intestinal lumen [luLI] and solving the models minimising the reaction.

### Differences in secretion fluxes between AD cases and controls

To investigate differences in reaction fluxes between cases and controls, we fitted linear regression using the fluxes as the response variable, the health status as an independent variable correcting for categorical APOE genotype, and sex being the WBM models sex-specific.

### Investigation of host-microbiota co-metabolism

To investigate the involvement of host-microbiota co-metabolism in formate production, we considered the difference in secretion of formate between the personalised and the germ-free models and we compared these differences with the community microbial secretion of the same metabolite to evaluate the amount of urine secretion directly correlated to the microbial secretion and the amount correlated to host-microbiota co-metabolism.

We confirmed the capacity of the whole-body models to produce formate from other microbial metabolites by adding individually to the diet metabolites that can be secreted by the microbiome. All metabolites were added to the diet setting their potential maximum intake to 300 mmol/d. These implemented diets were tested for maximum urine formate production on the germ-free models to evaluate changes in the previous baseline value.

### Genetic involvement in host formate production pathways in AD

We considered all the reactions known from the literature<sup>13</sup> to be involved in the formate metabolism including the ones we found responsible for the host-microbiota co-metabolism. We retrieved all the genes associated with these reactions from the Virtual Metabolic Human database (VMH)<sup>69</sup>, which houses the generic human metabolic reconstruction, Recon3D<sup>52</sup>, and then we investigated for associations between genes and health status using the AD Atlas<sup>35</sup> a network-based integrative resource for AD. The AD Atlas integrates multi-omics data from large-scale population-based and AD case-control studies, enabling users to annotate genes, metabolites, or phenotypes of interest in an AD-related context. In particular, differential expression genetic (DEG) data had been integrated in AD atlas from 2,114 postmortem human brain samples<sup>72</sup>. In our study, the AD atlas was queried using the genes identified to be involved in formate metabolism, retrieving information on brain-region-specific differential gene expression and differential protein abundance in AD, as well as genetic associations to AD-related phenotypes inferred from large-scale genome-wide association studies (GWAS). A gene-wise significance threshold was applied to filter GWAS results, as implemented by the AD Atlas<sup>35</sup>.

## Data availability

The data supporting the conclusions of this study are accessible; however, constraints are imposed on their public availability. Notwithstanding, these data can be procured from the authors upon the submission of a reasonable request. Data from the DELCODE cohort will be made available on request to Johannes Hertel (Johannes.Hertel@med.uni-greifswald.de). Data from the WISCONSIN cohort will be made available on request to Barbara B. Bendlin (bbb@medicine.wisc.edu).

Received: 29 August 2023; Accepted: 29 February 2024

Published online: 13 March 2024

## References

1. Ricci, M., Cimini, A., Chiaravallotti, A., Filippi, L. & Schillaci, O. Positron emission tomography (PET) and neuroimaging in the personalized approach to neurodegenerative causes of dementia. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms21207481> (2020).

2. Knopman, D. S. *et al.* Alzheimer disease. *Nat. Rev. Dis. Primers* **7**, 33. <https://doi.org/10.1038/s41572-021-00269-y> (2021).
3. Castro, D. M., Dillon, C., Machnicki, G. & Allegri, R. F. The economic cost of Alzheimer's disease: Family or public health burden?. *Dement. Neuropsychol.* **4**, 262–267. <https://doi.org/10.1590/S1980-57642010DN40400003> (2010).
4. Toledo, J. B. *et al.* Metabolic network failures in Alzheimer's disease: A biochemical road map. *Alzheimers Dement.* **13**, 965–984. <https://doi.org/10.1016/j.jalz.2017.01.020> (2017).
5. Lynch, S. V. & Pedersen, O. The human intestinal microbiome in health and disease. *N. Engl. J. Med.* **375**, 2369–2379. <https://doi.org/10.1056/NEJMra1600266> (2016).
6. Alexander, M. & Turnbaugh, P. J. Deconstructing mechanisms of diet-microbiome-immune interactions. *Immunity* **53**, 264–276. <https://doi.org/10.1016/j.immuni.2020.07.015> (2020).
7. Cryan, J. F. *et al.* The microbiota-gut-brain axis. *Physiol. Rev.* **99**, 1877–2013. <https://doi.org/10.1152/physrev.00018.2018> (2019).
8. Gheorghe, C. E. *et al.* Focus on the essentials: Tryptophan metabolism and the microbiome-gut-brain axis. *Curr. Opin. Pharmacol.* **48**, 137–145. <https://doi.org/10.1016/j.coph.2019.08.004> (2019).
9. Bonfili, L. *et al.* Microbiota modulation as preventative and therapeutic approach in Alzheimer's disease. *FEBS J.* <https://doi.org/10.1111/febs.15571> (2020).
10. Goyal, D., Ali, S. A. & Singh, R. K. Emerging role of gut microbiota in modulation of neuroinflammation and neurodegeneration with emphasis on Alzheimer's disease. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **106**, 110112. <https://doi.org/10.1016/j.pnpbb.2020.110112> (2021).
11. 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. <https://doi.org/10.1016/j.jalz.2018.07.217> (2019).
12. 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. <https://doi.org/10.1016/j.jalz.2018.08.012> (2019).
13. Pietzke, M., Meiser, J. & Vazquez, A. Formate metabolism in health and disease. *Mol. Metab.* **33**, 23–37. <https://doi.org/10.1016/j.molmet.2019.05.012> (2020).
14. Wang, Y. *et al.* Systematic evaluation of urinary formic acid as a new potential biomarker for Alzheimer's disease. *Front. Aging Neurosci.* <https://doi.org/10.3389/fnagi.2022.1046066> (2022).
15. Pálsson, B. Ø. *Systems Biology: Constraint-based Reconstruction and Analysis* (Cambridge University Press, 2015).
16. Thiele, I. & Pálsson, B. Ø. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protocols* **5**, 93–121 (2010).
17. Orth, J. D., Thiele, I. & Pálsson, B. O. What is flux balance analysis?. *Nat. Biotechnol.* **28**, 245–248 (2010).
18. Magnusdóttir, S. *et al.* Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat. Biotechnol.* **35**, 81–89. <https://doi.org/10.1038/nbt.3703> (2017).
19. Heinken, A. *et al.* Genome-scale metabolic reconstruction of 7,302 human microorganisms for personalized medicine. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-022-01628-0> (2023).
20. Heinken, A. *et al.* APOLLO: A genome-scale metabolic reconstruction resource of 247,092 diverse human microbes spanning multiple continents, age groups, and body sites. *bioRxiv* <https://doi.org/10.1101/2023.10.02.560573> (2023).
21. Baldini, F. *et al.* The microbiome modeling toolbox: From microbial interactions to personalized microbial communities. *Bioinformatics* <https://doi.org/10.1093/bioinformatics/bty941> (2018).
22. Heinken, A. & Thiele, I. Microbiome Modelling Toolbox 2.0: Efficient, tractable modelling of microbiome communities. *Bioinformatics* **38**, 2367–2368. <https://doi.org/10.1093/bioinformatics/btac082> (2022).
23. Hertel, J., Heinken, A., Martinelli, F. & Thiele, I. Integration of constraint-based modeling with fecal metabolomics reveals large deleterious effects of *Fusobacterium* spp. on community butyrate production. *Gut Microbes* **13**, 1–23. <https://doi.org/10.1080/19490976.2021.1915673> (2021).
24. Thiele, I. *et al.* Personalized whole-body models integrate metabolism, physiology, and the gut microbiome. *Mol. Syst. Biol.* **16**, e8982. <https://doi.org/10.15252/msb.20198982> (2020).
25. Basile, A. *et al.* Longitudinal flux balance analyses of a patient with episodic colonic inflammation reveals microbiome metabolic dynamics. *Gut Microbes* **15**, 2226921. <https://doi.org/10.1080/19490976.2023.2226921> (2023).
26. Thiele, I. & Fleming, R. M. T. Whole-body metabolic modelling predicts isoleucine dependency of SARS-CoV-2 replication. *Comput. Struct. Biotechnol. J.* **20**, 4098–4109. <https://doi.org/10.1016/j.csbj.2022.07.019> (2022).
27. Jessen, F. *et al.* Design and first baseline data of the DZNE multicenter observational study on predementia Alzheimer's disease (DELCODE). *Alzheimers Res. Ther.* **10**, 15. <https://doi.org/10.1186/s13195-017-0314-2> (2018).
28. Vogt, N. M. *et al.* Gut microbiome alterations in Alzheimer's disease. *Sci. Rep.* **7**, 13537. <https://doi.org/10.1038/s41598-017-13601-y> (2017).
29. Saunders, A. M. *et al.* Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* **43**, 1467–1472. <https://doi.org/10.1212/wnl.43.8.1467> (1993).
30. Naj, A. C. *et al.* Effects of multiple genetic loci on age at onset in late-onset Alzheimer disease: A genome-wide association study. *JAMA Neurol.* **71**, 1394–1404. <https://doi.org/10.1001/jamaneurol.2014.1491> (2014).
31. Zhu, Q. *et al.* Phylogeny-aware analysis of metagenome community ecology based on matched reference genomes while bypassing taxonomy. *mSystems* **7**, e00167–e122. <https://doi.org/10.1128/mSystems.00167-22> (2022).
32. Zhu, Q. *et al.* Phylogenomics of 10,575 genomes reveals evolutionary proximity between domains Bacteria and Archaea. *Nat. Commun.* **10**, 5477. <https://doi.org/10.1038/s41467-019-13443-4> (2019).
33. Gonzalez, A. *et al.* Qiita: Rapid, web-enabled microbiome meta-analysis. *Nat. Methods* **15**, 796–798. <https://doi.org/10.1038/s41592-018-0141-9> (2018).
34. Brosnan, M. E. & Brosnan, J. T. Formate: The neglected member of one-carbon metabolism. *Annu. Rev. Nutr.* **36**, 369–388. <https://doi.org/10.1146/annurev-nutr-071715-050738> (2016).
35. Wörheide, M. A., Krumsiek, J., Kastenmüller, G. & Arnold, M. Multi-omics integration in biomedical research: A metabolomics-centric review. *Anal. Chim. Acta* **1141**, 144–162. <https://doi.org/10.1016/j.aca.2020.10.038> (2021).
36. Watanabe, Y. *et al.* Alterations in glycerolipid and fatty acid metabolic pathways in Alzheimer's disease identified by urinary metabolic profiling: A pilot study. *Front. Neurol.* <https://doi.org/10.3389/fneur.2021.719159> (2021).
37. Yilmaz, A. *et al.* Targeted metabolic profiling of urine highlights a potential biomarker panel for the diagnosis of Alzheimer's disease and mild cognitive impairment: A pilot study. *Metabolites* <https://doi.org/10.3390/metabo10090357> (2020).
38. Boeniger, M. F. Formate in urine as a biological indicator of formaldehyde exposure: A review. *Am. Ind. Hyg. Assoc. J.* **48**, 900–908. <https://doi.org/10.1080/15298668791385787> (1987).
39. Hajjar, I., Liu, C., Jones, D. P. & Uppal, K. Untargeted metabolomics reveal dysregulations in sugar, methionine, and tyrosine pathways in the prodromal state of AD. *Alzheimer's Dement.* **12**, e12064. <https://doi.org/10.1002/dad2.12064> (2020).
40. Kaddurah-Daouk, R. *et al.* Alterations in metabolic pathways and networks in Alzheimer's disease. *Transl. Psychiatry* **3**, e244. <https://doi.org/10.1038/tp.2013.18> (2013).
41. Maitre, M., Klein, C., Patte-Mensah, C. & Mensah-Nyagan, A. G. Tryptophan metabolites modify brain A $\beta$  peptide degradation: A role in Alzheimer's disease?. *Prog. Neurobiol.* **190**, 101800. <https://doi.org/10.1016/j.pneurobio.2020.101800> (2020).
42. Clarke, J. R., Ribeiro, F. C., Frozza, R. L., De Felice, F. G. & Lourenco, M. V. Metabolic dysfunction in Alzheimer's disease: From basic neurobiology to clinical approaches. *J. Alzheimer's Dis.* **64**, S405–S426. <https://doi.org/10.3233/jad-179911> (2018).

43. Griffin, J. W. & Bradshaw, P. C. Amino acid catabolism in Alzheimer's disease brain: Friend or foe?. *Oxid. Med. Cell. Longev.* **2017**, 5472792. <https://doi.org/10.1155/2017/5472792> (2017).
44. Schwarcz, R. & Stone, T. W. The kynurenine pathway and the brain: Challenges, controversies and promises. *Neuropharmacology* **112**, 237–247. <https://doi.org/10.1016/j.neuropharm.2016.08.003> (2017).
45. Porter, R. J. *et al.* Cognitive deficit induced by acute tryptophan depletion in patients with Alzheimer's disease. *Am. J. Psychiatry* **157**, 638–640. <https://doi.org/10.1176/appi.ajp.157.4.638> (2000).
46. 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. *Alzheimers Res. Ther.* **13**, 20. <https://doi.org/10.1186/s13195-020-00741-z> (2021).
47. van der Velpen, V. *et al.* Systemic and central nervous system metabolic alterations in Alzheimer's disease. *Alzheimer's Res. Ther.* **11**, 93. <https://doi.org/10.1186/s13195-019-0551-7> (2019).
48. Tait-Mulder, J., Hodge, K., Sumpton, D., Zanivan, S. & Vazquez, A. The conversion of formate into purines stimulates mTORC1 leading to CAD-dependent activation of pyrimidine synthesis. *Cancer Metab.* **8**, 20. <https://doi.org/10.1186/s40170-020-00228-3> (2020).
49. Zhang, X. *et al.* The association between folate and Alzheimer's disease: A systematic review and meta-analysis. *Front. Neurosci.* <https://doi.org/10.3389/fnins.2021.661198> (2021).
50. Bergau, N., Maul, S., Rujescu, D., Simm, A. & Navarrete Santos, A. Reduction of glycolysis intermediate concentrations in the cerebrospinal fluid of Alzheimer's disease patients. *Front. Neurosci.* **13**, 871. <https://doi.org/10.3389/fnins.2019.00871> (2019).
51. Gu, C., Kim, G. B., Kim, W. J., Kim, H. U. & Lee, S. Y. Current status and applications of genome-scale metabolic models. *Genome Biol.* **20**, 121. <https://doi.org/10.1186/s13059-019-1730-3> (2019).
52. Brunk, E. *et al.* Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nat. Biotechnol.* **36**, 272–281. <https://doi.org/10.1038/nbt.4072> (2018).
53. Thiele, I. *et al.* A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* **31**, 419–425. <https://doi.org/10.1038/nbt.2488> (2013).
54. Heinken, A., Magnúsdóttir, S., Fleming, R. M. T. & Thiele, I. DEMETER: Efficient simultaneous curation of genome-scale reconstructions guided by experimental data and refined gene annotations. *Bioinformatics* **37**, 3974–3975. <https://doi.org/10.1093/bioinformatics/btab622> (2021).
55. Reimer, L. C. *et al.* BacDive in 2019: Bacterial phenotypic data for High-throughput biodiversity analysis. *Nucleic Acids Res.* **47**, D631–D636. <https://doi.org/10.1093/nar/gky879> (2018).
56. Jessen, F. *et al.* A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimers Dement.* **10**, 844–852. <https://doi.org/10.1016/j.jalz.2014.01.001> (2014).
57. Albert, M. S. *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. *Alzheimers Dement.* **7**, 270–279. <https://doi.org/10.1016/j.jalz.2011.03.008> (2011).
58. McKhann, G. M. *et al.* The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* **7**, 263–269. <https://doi.org/10.1016/j.jalz.2011.03.005> (2011).
59. Pietzner, M. *et al.* Urine metabolomics by (1)H-NMR spectroscopy indicates associations between serum 3,5-T2 concentrations and intermediary metabolism in euthyroid humans. *Eur. Thyroid J.* **4**, 92–100. <https://doi.org/10.1159/000381308> (2015).
60. Pietzner, M. *et al.* Hepatic steatosis is associated with adverse molecular signatures in subjects without diabetes. *J. Clin. Endocrinol. Metab.* **103**, 3856–3868. <https://doi.org/10.1210/je.2018-00999> (2018).
61. Hertel, J. *et al.* Dilution correction for dynamically influenced urinary analyte data. *Anal. Chim. Acta* **1032**, 18–31. <https://doi.org/10.1016/j.aca.2018.07.068> (2018).
62. Besser, L. *et al.* Version 3 of the National Alzheimer's Coordinating Center's uniform data set. *Alzheimer Dis. Assoc. Disord.* **32**, 351–358. <https://doi.org/10.1097/wad.0000000000000279> (2018).
63. Johnson, S. C. *et al.* The Wisconsin Registry for Alzheimer's Prevention: A review of findings and current directions. *Alzheimer's Dement.* **10**, 130–142. <https://doi.org/10.1016/j.dadm.2017.11.007> (2018).
64. Chen, S., Zhou, Y., Chen, Y. & Gu, J. FASTP: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890. <https://doi.org/10.1093/bioinformatics/bty560> (2018).
65. Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100. <https://doi.org/10.1093/bioinformatics/bty191> (2018).
66. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359. <https://doi.org/10.1038/nmeth.1923> (2012).
67. Heinken, A., Sahoo, S., Fleming, R. M. & Thiele, I. Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes* **4**, 28–40. <https://doi.org/10.4161/gmic.22370> (2013).
68. Thiele, I., Fleming, R. M., Bordbar, A., Schellenberger, J. & Palsson, B. Functional characterization of alternate optimal solutions of *Escherichia coli*'s transcriptional and translational machinery. *Biophys. J.* **98**, 2072–2081. <https://doi.org/10.1016/j.bpj.2010.01.060> (2010).
69. Noronha, A. *et al.* The Virtual Metabolic Human database: Integrating human and gut microbiome metabolism with nutrition and disease. *Nucleic Acids Res.* **47**, D614–D624. <https://doi.org/10.1093/nar/gky992> (2019).
70. Heirendt, L. *et al.* Creation and analysis of biochemical constraint-based models using the COBRA Toolbox vol 3.0. *Nat. Protocols* **14**, 639–702. <https://doi.org/10.1038/s41596-018-0098-2> (2019).
71. Gudmundsson, S. & Thiele, I. Computationally efficient flux variability analysis. *BMC Bioinform.* **11**, 489 (2010).
72. Wan, Y. W. *et al.* Meta-analysis of the Alzheimer's disease human brain transcriptome and functional dissection in mouse models. *Cell Rep.* **32**, 107908. <https://doi.org/10.1016/j.celrep.2020.107908> (2020).

## Acknowledgements

This study was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (757922) to IT, the Science Foundation Ireland under Grant number 12/RC/2273-P2. Funding for this project was also provided through NIA grant U19AG063744 Alzheimer's Gut Microbiome Project (AGMP), PI Kaddurah-Daouk at Duke University along with several academic institutions. Additional funding for the AD Atlas includes NIA R01AG046171, RF1AG051550, RF1AG057452, RF1AG059093, RF1AG058942, U01AG061359, U19AG063744 and FNIH: #DAOU16AMPA awarded to Dr Kaddurah-Daouk at Duke University for the in partnership with a large number of academic institutions. Funding for the Wisconsin studies includes a Wisconsin Partnership Program grant (BBB, FER), National Institute on Aging grant R01AG070973 (BBB, FER), National Institute on Aging grant P30AG062715 (SA), National Institute on Aging grant RF1AG027161 (SCJ), National Institute of General Medical Sciences grant T32GM081061 (MBH), and National Institute of Neurological Disorders And Stroke grant F99NS130922 (MBH). Additional funding was

provided by the National Institute on Aging (RF1AG058942, R01AG037639, R01AG021155, R01AG069901, R01AG054047, U01AG061359).

### Author contributions

I.T. and J.H. conceived the study. S.T. collected biosamples of the DELCODE cohort and helped interpret findings from the clinical data. Collaborators from the Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE) recruited participants and collected urine samples for the DELCODE cohort. S.A., C.D.E., M.H., S.C.J., F.E.R., B.B.B. recruited participants, collected faecal samples and provided data on age, sex, BMI, *APOE* genotype, diagnosis, and Trail Making Test B Score to the Alzheimer's Disease Metabolomics Consortium (ADMC). The Knight Lab, UCSD, provided the WGS data. F.M. prepared the relative abundances table, and F.M. and A.H. constructed the personalised microbiome models. F.M. performed the microbiome model simulations. J.H. and F.M. designed the statistical analyses. J.H. conducted the DELCODE analysis. M.A.W., M.A., and G.K. conducted the analyses based on the AD Atlas. A.K.H., K.B. and M.N. provided the NMR quantifications for the DELCODE study. H.J.G. supervised the DELCODE-associated analyses and calculations. F.M. and T.H. performed the reactions KO simulations. F.M. and I.T. performed the WBM simulations. I.T. devised and supervised the project. I.T., F.M., J.H., and A.H. wrote the initial draft manuscript. All authors discussed the results and contributed to the writing of the final manuscript. The investigators within the AGMP and the Alzheimer's Disease Metabolomics Consortium, not listed specifically in this publication's author's list, provided data along with its pre-processing and prepared it for analysis, but did not participate in the analysis or writing of this manuscript. A listing of AGMP Investigators can be found at <https://alzheimergut.org/meet-the-team/>. A complete listing of ADCM investigators can be found at: <https://sites.duke.edu/adnimetab/team/>. We are also grateful to the research participants in the Wisconsin Registry for Alzheimer's Prevention and the Wisconsin Alzheimer's Disease Research Center.

### Competing interests

G.K. and M.A. are co-inventors of several patent applications on the use of metabolomics in Alzheimer's disease and own equity and IP in Chymia LLC and IP in PsyProtix unrelated to this work. R.K.D. is an inventor of a series of patents on the use of metabolomics for the diagnosis and treatment of CNS diseases and holds equity in Metabolon Inc., Chymia LLC and PsyProtix. S.T. has served on national and international advisory boards of Roche, Eisai, Grifols, and Biogen, and is a member of the independent data safety and monitoring board of the study ENVISION (Biogen). H.J.G. has received travel grants and speaker's honoraria from Fresenius Medical Care, Neuraxpharm, Servier, and Janssen Cilag as well as research funding from Fresenius Medical Care. R.K. is a scientific advisory board member, and consultant for BiomeSense, Inc., has equity and receives income. He is a scientific advisory board member and has equity in GenCirq. He is a consultant and scientific advisory board member for DayTwo, and receives income. He has equity in and acts as a consultant for Cybele. He is a co-founder of Biota, Inc., and has equity. He is a cofounder of Micronoma, and has equity and is a scientific advisory board member. D.M. is a consultant for, and has stock in, BiomeSense, Inc. The terms of these arrangements have been reviewed and approved by the University of California, San Diego in accordance with its conflict of interest policies.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-55960-3>.

**Correspondence** and requests for materials should be addressed to J.H. or I.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024