Age-related microbiome metabolites modulate splicing and chromatin accessibility in the brain

Meenakshi Chakraborty[1], Sophia M. Shi[2,3,4,5], Imani E. Porter[1], Daniel J. Richard[1], Georgi K. Marinov[1], Ashley A. Moore[6], Jenna L. E. Blum[7], Aravind Natarajan[6], James W. Jahng[8,9], Joseph C. Wu[8,9], Sydney X. Lu[6], Shawn M. Davidson[7], William J. Greenleaf[1], Nay L. Saw[10], Mehrdad Shamloo[10,11], Anne Brunet[1,5,12,13], Tony Wyss-Coray[4,5,13], Ami S. Bhatt[1,6,13,*]

[1]Department of Genetics, Stanford University, Stanford, CA, USA. [2]Department of Chemistry, Stanford University, Stanford, CA, USA. [3]Stanford Chemistry, Engineering and Medicine for Human Health (ChEM-H), Stanford University, Stanford, CA, USA. [4]Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA, USA. [5]Wu Tsai Neurosciences Institute, Stanford University School of Medicine, Stanford, CA, USA. [6]Department of Medicine, Divisions of Hematology and Blood & Marrow Transplantation, Stanford University, Stanford, CA, USA. [7]Division of Pulmonary and Critical Care Medicine, Department of Medicine, The Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. [8]Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA, USA. [9]Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, USA. [10]Stanford Behavioral and Functional Neuroscience Laboratory, Wu Tsai Neurosciences Institute, Stanford University, Stanford, CA, USA. [11]Department of Neurosurgery, School of Medicine, Stanford University, Stanford, CA, USA. [12]Glenn Center for the Biology of Aging, Stanford University, Stanford, CA, USA. [13]The Phil and Penny Knight Initiative for Brain Resilience, Stanford University, Stanford, CA, USA.

*Correspondence should be addressed to Ami S. Bhatt (asbhatt@stanford.edu).

Abstract

The gut microbiome generates diverse metabolites that can enter the bloodstream and alter host biology, including brain function. Hundreds of physiologically relevant, gut-brain signaling molecules likely exist; however, there has been no systematic, high-throughput effort to identify and validate them. Here, we integrate computational, *in vitro*, and *in vivo* approaches to pinpoint microbiome-derived metabolites whose blood levels change during aging, and that induce corresponding changes in the mouse brain. First, we mine large-scale metabolomics datasets from human cohorts (each $n \geq 1200$) to identify 30 microbiome-associated metabolites whose blood levels change with age. We then screen this panel in an *in vitro* transcriptomic assay to identify metabolites that perturb genes linked to age-related neurodegeneration. We then test four metabolites in an acute-exposure mouse model, and use multi-omic approaches to evaluate their impact on cellular functions in the brain. We confirm the known neurodegeneration-promoting effects of trimethylamine N-oxide (TMAO), including mitochondrial dysfunction, and further discover its disruptive impact on the pathways of glycolysis, GABAergic signaling, and RNA splicing. Additionally, we identify glycodeoxycholic acid (GDCA), a microbiome-derived secondary bile acid, as a potent regulator of chromatin accessibility and suppressor of genes that protect the brain from age-related, neurodegeneration-promoting insults. GDCA also acutely reduces mobility. Taken together, this work identifies microbiome-derived signals relevant to age-related neurodegeneration, and defines a scalable framework for linking microbiome metabolites to host pathologies.

Introduction

Aging is the greatest risk factor for many neurodegenerative diseases. Remarkably, age-related conditions that often precede these diseases, such as mild cognitive impairment (MCI), can be reversible[1,2]. Thus, identifying potentially modifiable age-related factors that promote neurodegeneration could alter and perhaps improve the natural history of disease initiation and progression.

One physiological compartment that changes with aging is the gut microbiome[3–5], likely due to factors such as increased permeability of the gut lining, increased medication usage, immunosenescence, and accumulated exposure to environmental toxins. In light of growing interest in the microbiota-gut-brain axis[6,7], a natural question arises: does “microbiome aging” contribute to “brain aging” and associated pathologies? Fecal microbiota transfers between young and aged mice - which can induce or reverse neuroinflammation and cognitive decline - suggest that it does[8–10]. However, the microbiome-related molecules and underlying molecular mechanisms that drive these phenotypes remain largely uncharacterized in both animals and humans, which limits translation of this knowledge to the clinic. This gap in the gut-brain field is partly due to the underutilization of *in vitro* systems, which can help create a pipeline from hypothesis-generating data analyses to *in vivo* validation.

The microbiome is known to substantially impact the levels of blood metabolites, with hundreds of metabolites differing in abundance between germ-free mice and their conventionally raised counterparts[11,12]. This is significant because blood metabolites enable the proper functioning of host organs; when their levels change, this can have profound effects on physiology[13–15]. We therefore hypothesized that microbiome aging affects brain aging by altering the levels of bioactive blood metabolites. These metabolites could impact the blood-brain barrier - a key regulator of brain homeostasis whose functionality declines with age[16] - and perhaps even cross it to act directly on neurons and glia.

To evaluate this hypothesis, we first leveraged public datasets[12,17–23] to identify a panel of 30 blood metabolites whose levels: 1) change with human aging, and 2) are significantly associated with the presence and/or composition of the gut microbiome. We then screened these candidates *in vitro* for their ability to alter the expression of genes in pathways related to age-related neurodegeneration, and validated four of the most interesting metabolites *in vivo* through multi-omic analyses of the mouse brain. Our results confirm and expand the neurodegeneration-promoting effects of the microbiome-derived metabolite trimethylamine N-oxide (TMAO), including novel effects on RNA splicing. In addition, we find that glycodeoxycholic acid (GDCA), a microbiome-derived bile acid, alters chromatin accessibility and silences many genes that maintain brain homeostasis after stress and injury. Finally, in behavioral experiments, we find that GDCA acutely reduces mobility. Together, our findings on TMAO and GDCA, our curated panel of age- and microbiome-associated blood metabolites, and the integrative framework we present - combining computational, *in vitro*, and *in vivo* approaches - provide a foundation for substantial advances in identifying microbiome-derived molecules that play key roles in host physiology.

Results

Identifying microbiome-associated metabolites whose blood levels change with age

There have been extensive efforts to characterize the specific metabolomic changes that occur with aging[17]. In addition, recent work has identified metabolites associated with microbiome composition and activity[11,12,21–23]. We leveraged these data to build a panel of metabolites for medium-throughput *in vitro* testing (Fig. 1; Supplementary Table 1; Methods).

Specifically, we first referenced a recent review on the metabolomics of aging[17] to identify three large cohort studies[18–20] that each analyzed blood samples from at least 1,200 participants. Pooling data across these studies yielded hundreds of metabolites whose blood levels changed with age in at least one study. Separately, we curated a list of human-associated metabolites that have been reported to be produced by microbes, either alone or via host-microbe "cometabolism" (where both host and microbial metabolism contribute to the final product). Intersecting this curated list with the age-associated metabolites yielded 79 unique overlapping metabolites.

For experimental feasibility, we further narrowed the list, prioritizing metabolites most strongly linked to the gut microbiome. We used two key resources: **(1)** a recent study comparing blood metabolite levels between germ-free and conventionally raised mice[12]; and **(2)** the GUTSY Atlas[22], which reported the proportion of variance in blood metabolite levels "explained" by the gut microbiome for 1,168 metabolites based on paired stool-blood data for over 8,000 humans. Metabolites were retained for the final panel if they met at least one of two criteria: **(1)** showed significantly different blood levels between germ-free and conventionally raised mice[12]; **(2)** ranked in the top 20% of blood metabolites for the GUTSY variance-explained metric.

We eliminated alanine because it is expected to have minimal impact on cellular processes[24], and uridine because its blood levels increased with age in one study[18] but decreased with age in another[20]. The final panel comprised 30 metabolites, 26 of which increased with age and 4 of which decreased with age (Supplementary Table 1). The metabolites encompassed a broad spectrum of categories, spanning six super classes, thirteen classes, and sixteen sub classes within the Human Metabolome Database (HMDB)[25] taxonomy (Fig. S1).

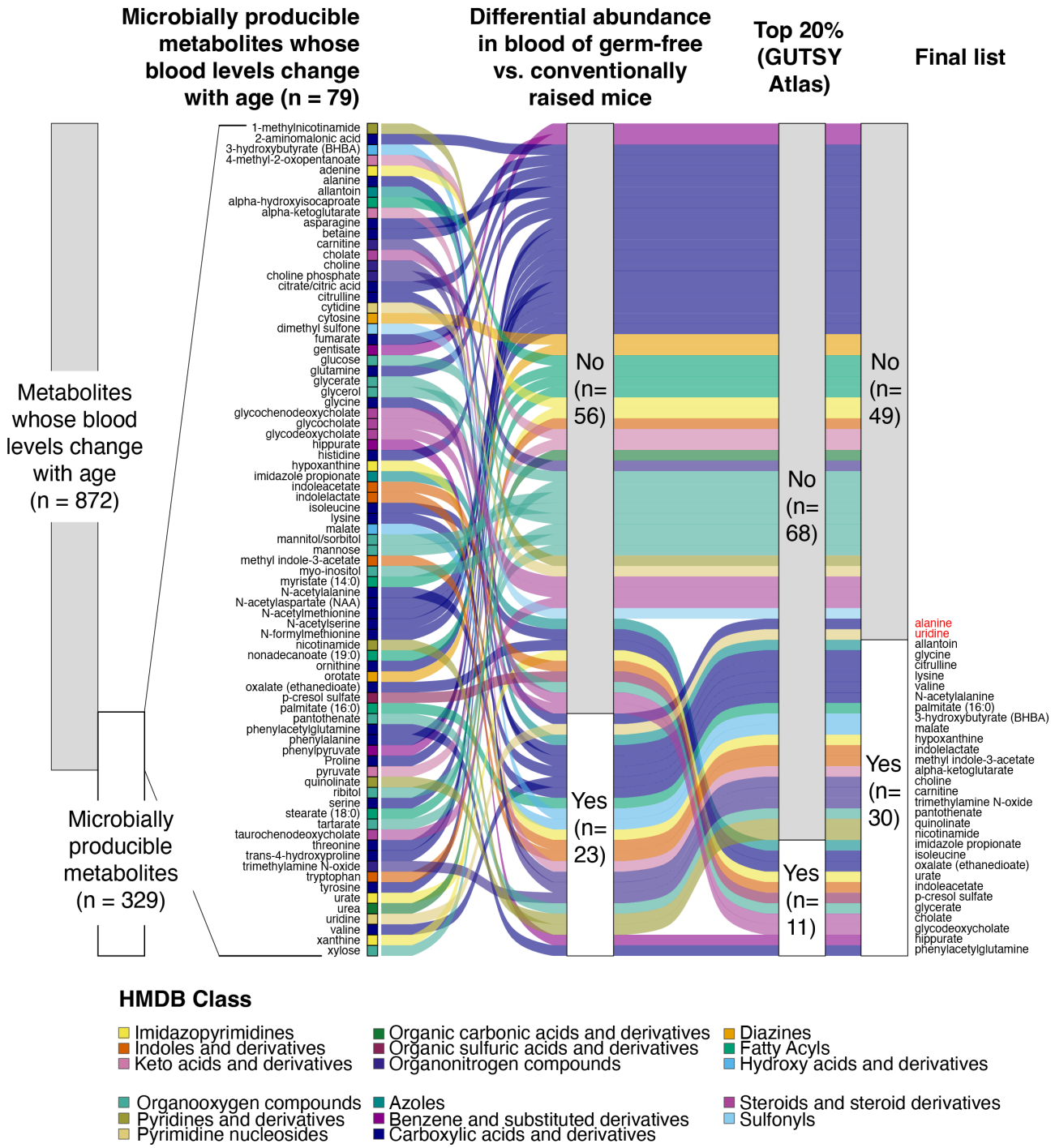


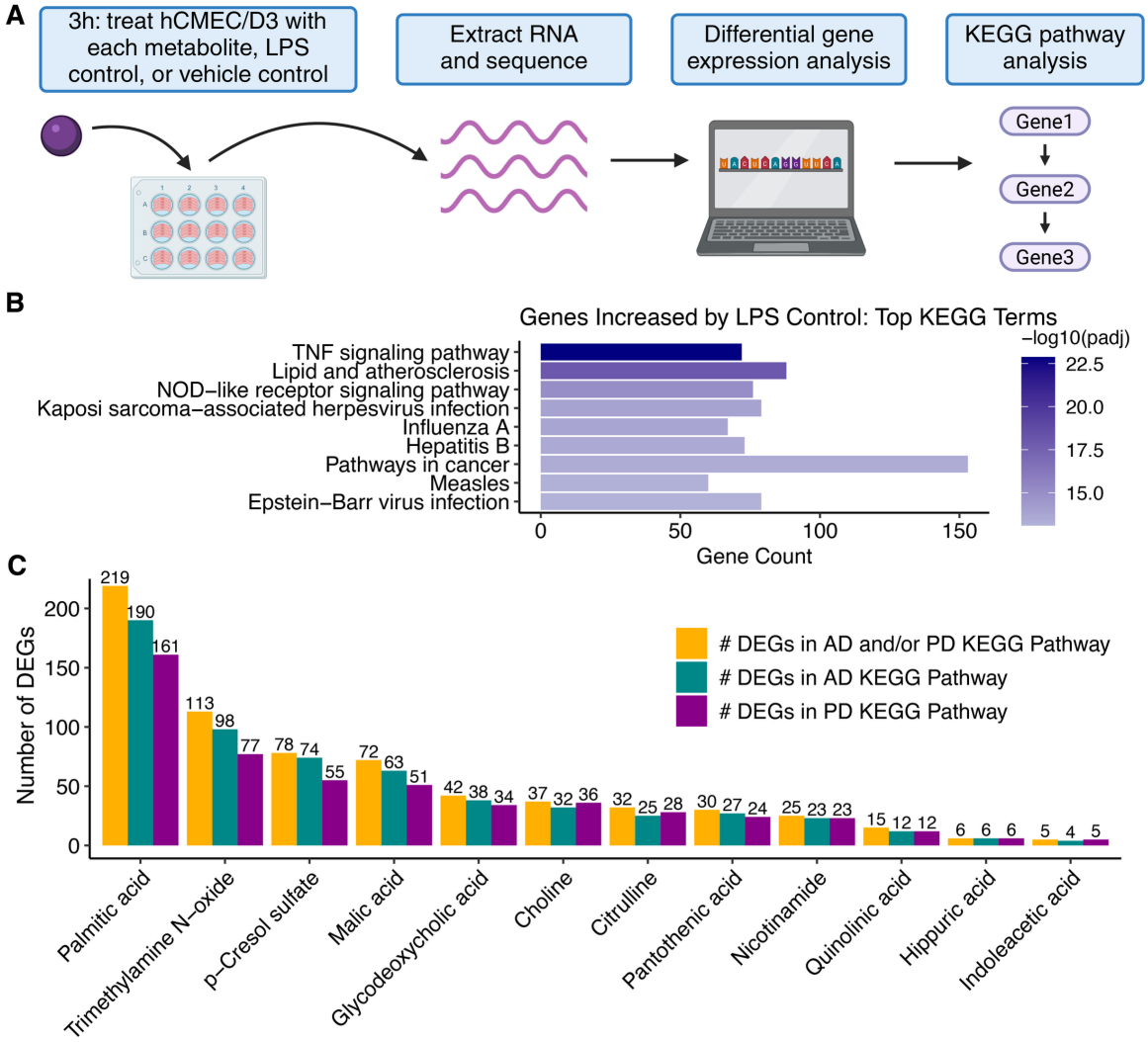
Fig. 1. Workflow used to define a panel of 30 microbiome-associated metabolites whose blood levels change with age. First, metabolites whose blood levels changed with age in large cohort(s) were overlapped with a curated list of metabolites that are producible by microbes, either alone or via host-microbe cometabolism. Metabolites were retained for the final panel if they met at least one of two criteria: **(1)** showed significantly different blood levels in germ-free vs. conventionally raised mice (based on Lai *et al.*, 2021[12]), and/or **(2)** ranked in the top 20% for the proportion of

124 variance in blood levels "explained" by the gut microbiome (metric from the GUTSY Atlas[22]). After
125 excluding alanine and uridine (see Results), the final panel comprised 30 metabolites.

126 *In vitro* screening reveals that microbiome-associated metabolites perturb 127 genes and pathways linked to age-related neurodegeneration

128 While *in vivo* validation remains the gold standard for assessing physiological relevance, we
129 began with *in vitro* screening, as testing dozens of metabolites *in vivo* for potential effects on the
130 brain would not have been practically feasible. We chose human brain endothelial cells as our
131 model system because they line the brain's blood vessels and are the first brain cell type
132 encountered by blood metabolites. In fact, many blood metabolites never reach other brain cell
133 types, since endothelial cells are a key component of the blood-brain barrier (BBB), which
134 selectively regulates entry into the brain parenchyma. Importantly, even metabolites that do not
135 cross the BBB could significantly impact brain function, since BBB decline is a defining feature of
136 brain aging and neurodegeneration[16].
137 We chose to use hCMEC/D3 cells, an immortalized human brain endothelial cell line
138 widely used in BBB studies[26]. To conduct the screen, each metabolite was added individually to
139 low-passage hCMEC/D3 cells in triplicate, alongside matched vehicle controls containing the
140 corresponding solvent in the media (Fig. 2A; Supplementary Table 3). After a 3-hour incubation
141 (chosen over longer timepoints to minimize metabolite clearance), cells were lysed in TRIzol for
142 downstream RNA extraction and sequencing. Each metabolite was added at a representative
143 physiological concentration, primarily based on its listed blood concentrations in HMDB (Methods;
144 Supplementary Table 1). In addition to vehicle controls, we included lipopolysaccharide (LPS)
145 controls in each sequencing batch, given the known effects of LPS on human brain endothelial
146 cells[27]. As expected, LPS induced the expression of genes in KEGG pathways associated with
147 inflammation, such as "TNF signaling pathway" and "NOD-like receptor signaling pathway" (Fig.
148 2B; Fig. S2; Supplementary Table 2).
149 The number of differentially expressed genes (DEGs), relative to vehicle controls, varied
150 considerably across the 30 test metabolites, reflecting substantial differences in the magnitude of
151 their transcriptional impact (Fig. S3; Supplementary Tables 4-5). The median number of DEGs
152 was 45.5, with an interquartile range (IQR) of 18.5 to 359.5, and a total range of 3 to 6,419 genes
153 affected. Twelve metabolites perturbed ≥5 genes in the KEGG pathways for "Alzheimer's
154 Disease" (AD) and/or "Parkinson's Disease" (PD), which are the two most common age-related
155 neurodegenerative disorders (Fig. 2C; Supplementary Table 6). The affected AD/PD genes
156 converged on processes including energy metabolism, proteasome-mediated protein
157 degradation, and inflammatory signaling, whose disruptions are central to neurodegenerative
158 disease[28].
159 We prioritized four metabolites for *in vivo* validation of their effects on brain function:
160 indoleacetic acid (IAA), glycodeoxycholic acid (GDCA), trimethylamine N-oxide (TMAO), and
161 palmitic acid (PA). TMAO's effects on the brain have been studied previously[29], so we included it
162 partially as a control and also to expand the scope of knowledge about its impact on the brain. All
163 four chosen metabolites share the following characteristics (Fig. 2D). They **(1)** impact the
164 expression of at least 5 unique genes in the AD and/or PD KEGG pathways in our transcriptomic
165 screen (Fig. 2C), **(2)** increase in the blood with age (Supplementary Table 1), **(3)** have been

associated with AD, PD, and/or impaired cognition in humans[29–39], **(4)** are strongly associated with production or regulation by the gut microbiome[12,22,40,41], and **(5)** have been detected in human cerebrospinal fluid or brain[42–46].



Criteria	IAA	GDCA	TMAO	PA
Unique AD/PD genes	5	42	113	219
Up with age?	Y	Y	Y	Y
Associated with AD, PD, or impaired cognition?	Y	Y	Y	Y
Strong association with gut microbiome?	Y: 92nd percentile GUTSY	Y: 91st percentile GUTSY	Y: near-absent in GF mice	Y: ~2x higher in GF mice
Found in CSF or brain?	Y	Y	Y	Y

Fig. 2. *In vitro* screening reveals that microbiome-associated metabolites perturb genes and pathways linked to age-associated neurodegenerative disease. **(A)** Schematic illustrating the

173 setup of the *in vitro* transcriptomic screen. The full panel of 30 metabolites was distributed across
174 four sequencing batches, with each metabolite included in one batch. Within each batch, at least
175 *n*=3 wells were treated with each assigned metabolite, its respective vehicle control, or the LPS
176 control (see Methods and Supplementary Table 3). **(B)** Top enriched KEGG pathways for the
177 genes significantly upregulated by the LPS control in the first sequencing batch. The analogous
178 results for batches #2-4 are presented in Fig. S2. The non-specific "KEGG root term" was filtered
179 out. **(C)** Number of differentially expressed genes (DEGs) per metabolite in the KEGG pathways
180 for Alzheimer's Disease (AD) and Parkinson's Disease (PD). Metabolites impacting at least 5
181 unique AD/PD genes are shown. **(D)** Identity and characteristics of the four metabolites chosen
182 for subsequent *in vivo* validation of their effects on the brain. IAA = indoleacetic acid, GDCA =
183 glycodeoxycholic acid, TMAO = trimethylamine N-oxide, PA = palmitic acid, AD = Alzheimer's
184 Disease, PD = Parkinson's Disease, GF = germ-free, CSF = cerebrospinal fluid. GUTSY refers to
185 the GUTSY Atlas[22], which reports microbiome-attributable variance in blood levels for 1,168 blood
186 metabolites. Structure diagrams were produced with rdkit[47]. Relevant studies (e.g., those
187 reporting metabolite detection in CSF or brain) are cited in the text.

188 TMAO disrupts transcriptional homeostasis and RNA splicing in the mouse 189 brain

190 To interrogate the effects of each of the four chosen metabolites on cellular functions in the brain
191 *in vivo*, we conducted bulk transcriptomics on mouse brain hemispheres following intravenous
192 retro-orbital injection of each metabolite or vehicle (2-hour treatment; *n*=4 male mice per injection
193 group; Fig. 3A). Doses were chosen to reflect physiologically relevant exposures, by converting
194 blood concentrations observed in seniors[32,34,48–55] to mouse-equivalent doses using well-
195 established scaling guidelines[56,57]. This experiment design was informed by prior work on
196 TMAO[58], such that TMAO could serve as a positive control. The main modification was the use
197 of intravenous retro-orbital injection in place of intraperitoneal injection, to more directly assess
198 the metabolite's impact on the brain.
199 For TMAO, the chosen mouse dose was 1.8 mg/kg, the same as the original study[58],
200 approximating 30μM concentration in human blood. The remaining doses were 22.2 mg/kg for PA
201 (~100μM human), 1.5 mg/kg for IAA (~10μM human), and 0.4 mg/kg for GDCA (~1μM human).
202 All human-equivalent concentrations were non-cytotoxic *in vitro* in hCMEC/D3 cells (Fig. S7). For
203 PA - unlike for the other metabolites - the human-equivalent concentration used for the mouse
204 experiment was reduced relative to the concentration used in the *in vitro* transcriptomic screen.
205 This was because the original concentration of 500μM was cytotoxic to hCMEC/D3 cells (Fig. S7),
206 likely explaining the disproportionately large number of DEGs observed for PA *in vitro* (Figs. 2C
207 and S3). Overall, the cytotoxicity testing (Fig. S7) ensured that the mouse exposures reflected
208 physiologically relevant human levels while avoiding human concentrations that are overtly
209 cytotoxic.
210 PA and IAA had negligible effects on gene expression in the mouse brain, with fewer than
211 ten differentially expressed genes per metabolite (Fig. 3B; Supplementary Table 7). By contrast,
212 TMAO and GDCA each had a significant transcriptional impact in the brain, modulating the
213 expression of 2,305 and 179 genes, respectively (Fig. 3C; Supplementary Table 7). We first

evaluated the TMAO results to verify concordance with the literature and also to derive additional insights about the impact of this microbiome metabolite on brain function.

In agreement with prior literature[58], we observed a coordinated downregulation of mitochondrially-encoded genes involved in oxidative phosphorylation, including *mt-Nd1, mt-Nd2, mt-Nd3, mt-Co2, mt-Co3, mt-Cytb, mt-Atp6,* and *mt-Atp8* (Supplementary Table 7). However, gene set enrichment analysis (GSEA) based on hallmark gene sets revealed a more complex rewiring of bioenergetic pathways (Fig. 4A; Supplementary Table 8). Despite the clear downregulation of mitochondrial oxidative phosphorylation, the term “oxidative phosphorylation” was positively enriched in GSEA when comparing TMAO-treated brains to controls. This result was due to the upregulation of several nuclear-encoded genes involved in oxidative phosphorylation (e.g., *Cox4i1, Cox6a1, Cox6b1, Cox7a2l, Ndufa7, Ndufa8, Ndufb2, Ndufb6, Ndufb8, Ndufc2*), perhaps reflecting a compensatory response, similar to that observed in patients with macro-deletion of mtDNA[59].

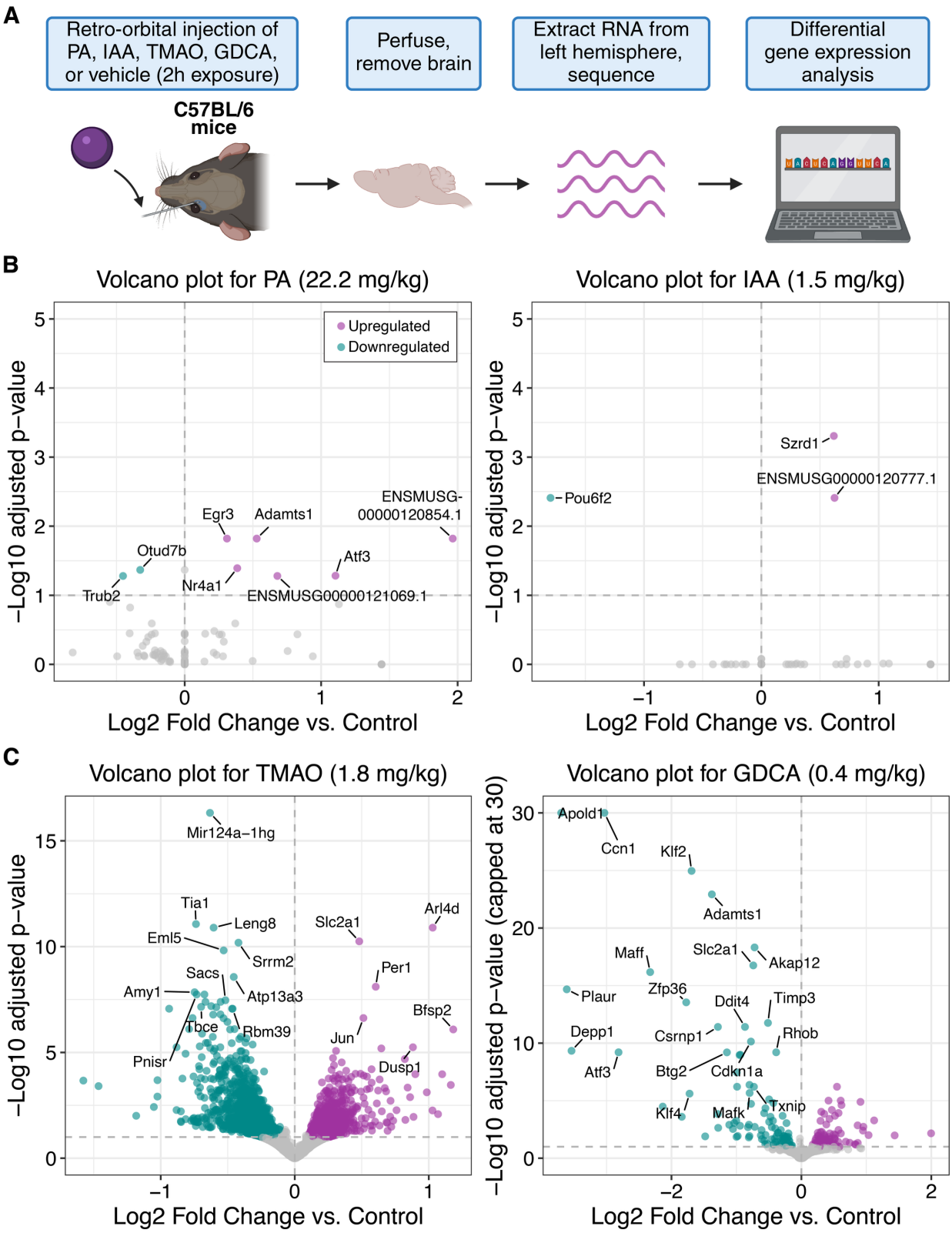


Fig. 3. The four metabolites selected for *in vivo* experiments impact brain gene expression to markedly different extents. **(A)** Schematic illustrating the setup of the *in vivo* transcriptomics

experiment; *n*=4 C57BL/6 male mice were injected with each metabolite or vehicle. **(B)** Volcano plots showing the results of differential gene expression analysis between PA- and IAA-treated brains and their vehicle controls. **(C)** Volcano plots showing the results of differential gene expression analysis between TMAO- and GDCA-treated brains and their vehicle controls. In both **(B)** and **(C)**, the y-axis represents the -log10(adjusted *p*-value) from DESeq2, and the x-axis represents the log2FoldChange vs. vehicle control, after applying the recommended log2FoldChange shrinkage via lfcShrink[60]. Purple dots represent significantly upregulated genes, with an adjusted $p \leq 0.1$ (DESeq2 default) and shrunken log2FoldChange ≥ 0.1. Teal dots represent significantly downregulated genes (adjusted $p \leq 0.1$ and shrunken log2FoldChange ≤ -0.1). The y-axis for the GDCA volcano plot was capped at 30 to avoid distortion by extreme values; the non-capped, adjusted *p*-values for *Apold1* and *Ccn1* are 1.82E-108 and 3.50E-70, respectively (Supplementary Table 7).

In addition, GSEA highlighted the term "glycolysis" as enriched (Fig. 4A). Indeed, eight glycolytic genes were upregulated by TMAO, including *Hk1* and *Pkm*, which catalyze rate-limiting steps of glycolysis[61] (Fig. 4B; Supplementary Table 7). Furthermore, *Slc2a1* (i.e., *Glut1*), a major glucose transporter in the brain that supports glycolytic flux, was the second most significantly upregulated gene by TMAO (adjusted *p*-value 5.62E-11; Fig. 3C and Supplementary Table 7). These findings agreed with data from our *in vitro* transcriptomic screen, where TMAO upregulated human glycolytic genes including *HK1*, *PKM*, *PFKP*, *GPI*, and *ALDOA* (Supplementary Table 5). Overall, our results align with previous findings that TMAO results in mitochondrial dysfunction in the brain[29]. Furthermore, our findings demonstrate a previously unreported upregulation of the glycolytic pathway in the brain.

Mitochondrial dysfunction can cause oxidative stress. Consistent with this, the "reactive oxygen species pathway" was significantly more highly expressed (Fig. 4A) in the brain transcriptome of TMAO-exposed animals vs. controls. In addition, *Txnip* - a known inducer of oxidative stress and activator of the inflammasome[62] - was significantly upregulated by TMAO (Supplementary Table 7). This finding is consistent with the most enriched GSEA term, "TNFα signaling via NF-κB" (Fig. 4A), which reflects inflammatory signaling. Together, these results support previous studies showing that TMAO promotes oxidative stress and inflammation, including in the brain[29,63,64].

GSEA using the REACTOME pathway database provided additional insights, including the positive enrichment of genes involved in synaptic transmission and GABAergic signaling, especially genes related to $GABA_B$ receptor activation (Fig. 4C; Supplementary Table 8). By contrast, we observed a marked downregulation of genes encoding $GABA_A$ receptors (Fig. S8). These results suggest that TMAO may shift inhibitory neural circuits from fast $GABA_A$-mediated responses towards slower, $GABA_B$-mediated responses. Although previous literature has demonstrated negative effects of TMAO on synaptic function[29], this is the first observation, to our knowledge, of specific perturbation of GABAergic signaling pathways. To assess the extent of TMAO's impact on neuronal cells vs. other brain cell types, we computed "marker genes" for six different brain cell classes, using a well-characterized single-cell dataset from the mouse brain[65] (Methods). TMAO downregulated over 15% of the marker genes for cells in the neuronal lineage, representing more than twice its impact on any of the other five cell classes (Fig. 4D).

GSEA also highlighted the depletion of genes associated with “mRNA splicing” (Fig. 4C). Strikingly, several of the genes that were most significantly downregulated by TMAO (e.g., *Tia1*, *Srrm2*, *Pnisr*, and *Rbm39*, see Fig. 3C and Supplementary Table 7) regulate alternative splicing[66–69]. This result was intriguing given the well-established link between changes in splicing patterns and the development of neurological disease[70].

Given the observed impact of TMAO on splicing-related factors, we sought to determine whether any genes were differentially spliced in TMAO-treated animals vs. controls. In agreement with the changes in splicing factor expression, the widely used rMATS-turbo software[71] nominated 126 differential splicing events between TMAO-treated and vehicle-treated brains, which we narrowed to 44 using conservative filters (Fig. 4E; Supplementary Table 9; Methods). Manual inspection in Integrative Genomics Viewer (IGV) revealed clear patterns consistent with differential splicing, including that of *Kmt2b*, a gene whose aberrant splicing has been linked to dystonia[72,73], a movement disorder that can occur independently or as a symptom of Parkinson’s disease (Fig. S9). *Kmt2b* also exhibited differential expression in TMAO- vs. vehicle-treated brains (log2FoldChange -0.39, adjusted *p* 0.001; Supplementary Table 7), potentially as a consequence of its differential splicing.

To our knowledge, this is the first demonstration of TMAO regulating alternative splicing patterns. Overall, our findings confirm[29,58] and broaden current understanding of TMAO’s neurodegeneration-promoting effects, which we summarize in a unified model (Fig. 4F).

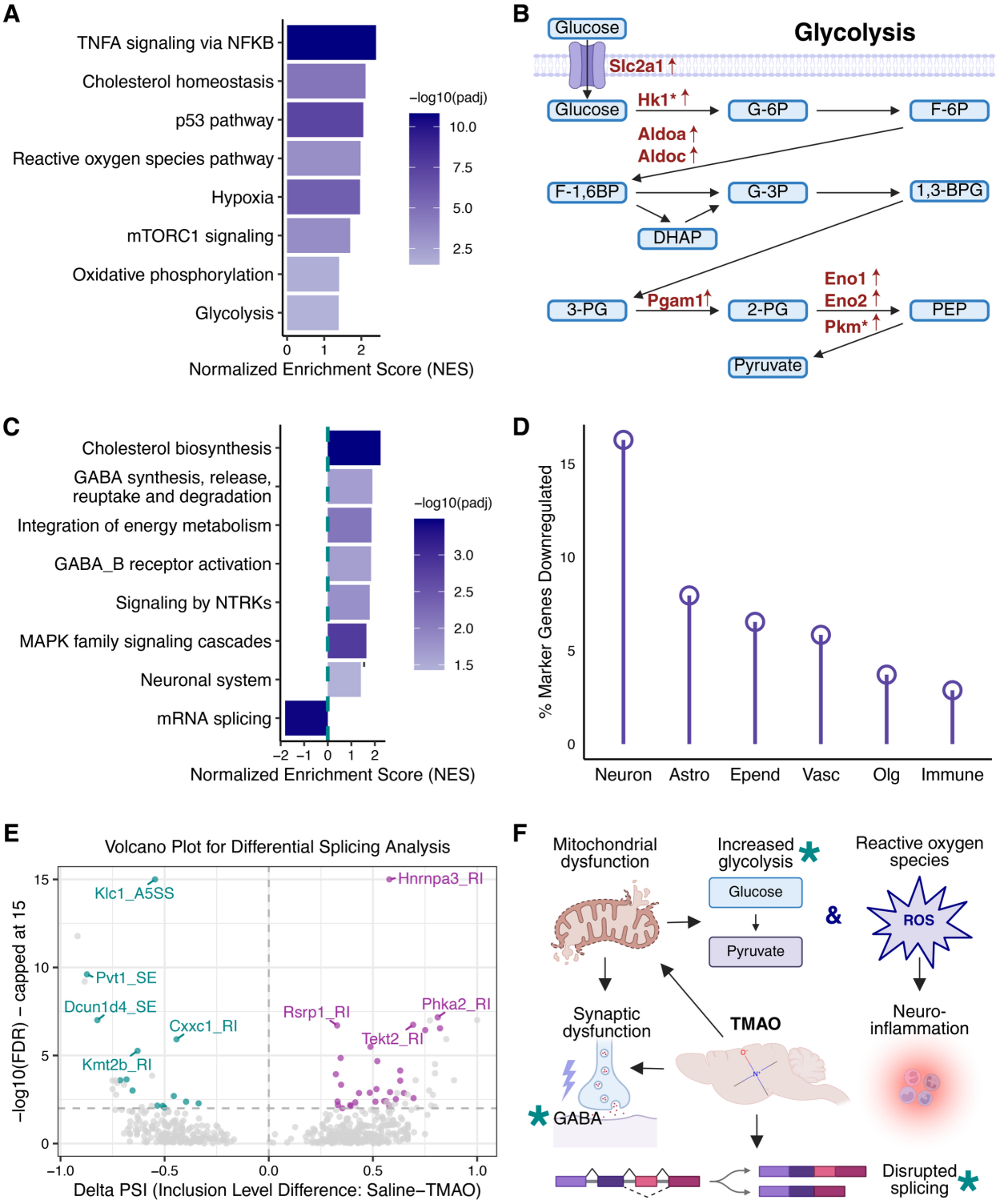


Fig. 4. TMAO disrupts transcriptional homeostasis and RNA splicing in the mouse brain. **(A)** Selected significantly enriched terms from gene set enrichment analysis (GSEA) based on hallmark gene sets; full results in Supplementary Table 8. **(B)** The glycolysis pathway, annotated with the glycolytic genes upregulated by TMAO. * = genes whose products catalyze rate-limiting

steps of glycolysis. Base figure adapted from Hu *et al.*, 2014[74]. **(C)** Selected significantly enriched terms from GSEA using the REACTOME pathway database; full results in Supplementary Table 8. **(D)** Percentage of marker genes in each brain cell class that are downregulated by TMAO. Neuron = neuronal lineage, Astro = astrocyte lineage and stem cells, Ependy = ependymal cells, Vasc = vasculature cells, Olg = oligodendrocyte lineage, Immune = immune cells. **(E)** Results of differential splicing analysis comparing brains from TMAO-exposed and control mice. Each point represents a splicing event. The x-axis represents the inclusion level difference (ΔPSI) and the y-axis represents the -log10(FDR), capped at 15 for visual clarity. Events significantly enriched in control or TMAO are shown in purple and teal, respectively; non-significant events are shown in gray. To reduce visual clutter, points were filtered to those with -log10(FDR) > 0. RI = retained intron, A5SS = alternative 5' splice site, SE = skipped exon. **(F)** Unified model of TMAO's neurodegeneration-promoting effects on the brain. Asterisks indicate portions of the model that are novel based on insights derived from this study. TMAO induces mitochondrial dysfunction, leading to compensatory upregulation of glycolysis and increased production of reactive oxygen species (ROS). The resulting oxidative stress further promotes neuroinflammation. Perhaps as a result of the rewired bioenergetic state, TMAO also promotes synaptic dysfunction, including impairments in GABAergic transmission. In addition, it downregulates splicing factors, resulting in changes to splicing patterns.

GDCA remodels chromatin and dramatically suppresses genes that protect the brain from stress and injury

Like TMAO, GDCA is a well-known microbiome-derived metabolite. It is a secondary conjugated bile acid, formed when the host conjugates microbially produced DCA with glycine. Blood levels of GDCA have been associated with neurodegeneration and/or cognitive impairment in at least five studies[30–34] (summarized in Supplementary Table 10). However, GDCA's brain-specific effects have not, to our knowledge, been experimentally evaluated in mammals.

GDCA strongly suppressed the expression of many genes in the mouse brain (Fig. 3C; Supplementary Table 7). The most significantly downregulated gene was *Apold1* (log2FoldChange -3.7, adjusted *p*-value 1.82E-108), a gene whose expression is critical for stroke recovery in mice[75]. *Ccn1*, the second most downregulated gene (log2FoldChange -3.0, adjusted *p*-value 3.50E-70), similarly promotes wound healing and tissue repair[76]. This pattern extended beyond individual genes: many of the genes that were strongly downregulated by GDCA (e.g., *Klf2*, *Klf4*, *Adamts1*, *Akap12*, *Maff*, *Plaur*, *Ddit4*, *Atf3*, *Btg2*) are involved in cellular responses to injury and stress, including oxidative stress, hypoxia, DNA damage, and stroke[77–85]. These observations were supported by GSEA based on hallmark gene sets. GSEA demonstrated significant negative enrichment of the terms "TNFα signaling via NF-κB" and "hypoxia" (Fig. 5A), as well as nominal negative enrichment of "p53 pathway" (Supplementary Table 8), consistent with reduced activity of stress-response pathways. The results suggested marked disruption of endothelial cell homeostasis; for example, the simultaneous reductions of *Klf2* and *Klf4* alone would be expected to cause severe endothelial dysfunction[86,87]. However, given the use of bulk RNA-seq, it is unknown whether these decreases originated specifically from endothelial cells, or from other brain cell types.

340 GSEA using the REACTOME pathway database (Fig. 5B; Supplementary Table 8)
341 demonstrated the positive enrichment of several terms related to chromatin dynamics, including
342 "PKMTs methylate histone lysines," "epigenetic regulation of gene expression," and "chromatin
343 organization." Based on this result, we postulated that GDCA achieves its notable gene regulatory
344 effects, including the strong silencing of many genes, by inducing changes in chromatin
345 organization and accessibility. To test this hypothesis, we conducted bulk ATAC-Seq on the flash-
346 frozen right brain hemispheres of the GDCA- and vehicle-treated mice.
347 Consistent with our prediction that GDCA induces chromatin changes, ATAC-seq
348 identified hundreds of genomic regions with differential accessibility between the brains of GDCA-
349 and vehicle-treated mice (Figs. 5C-D; Figs. S10A-B; Supplementary Table 11). 200 of the 204
350 differentially accessible peaks (98%) were less accessible in the GDCA-exposed brains.
351 Moreover, using HOMER to annotate each differential peak with its nearest gene, we found that
352 six of the "nearest genes" were also differentially expressed in RNA-seq (Supplementary Table
353 11). This set of six genes included several known stress-response genes: *Ddit4*[82], *Errfi1*[88], *Rhob*[89],
354 and *Sgk1*[90]. For all six genes, the direction of change was concordant across datasets - decreased
355 expression of the gene and decreased accessibility of the nearby peak - providing compelling
356 evidence that the strong transcriptional changes induced by GDCA are driven, at least in part, by
357 chromatin remodeling.
358 We further annotated the differentially accessible peaks based on overlapping genomic
359 features and compared the proportion of peaks in each category to the proportions expected from
360 the reference genome. This analysis revealed enrichment of differential peaks in promoters and
361 introns (Fig. 5E; Fig. S10C; Supplementary Table 11). Promoters and introns frequently harbor
362 regulatory elements such as enhancers[91]; therefore, the enrichment of these features among
363 differential peaks further supports a direct link between GDCA-induced changes in chromatin
364 accessibility and effects on gene expression. Consistent with this interpretation, 4 of the 5 most
365 significantly differential intronic peaks fully overlapped elements annotated as distal enhancers in
366 the ENCODE SCREEN database (https://screen.wenglab.org/)[92].
367 Promoters and enhancers commonly contain motifs that bind transcription factors (TFs)
368 critical for the regulation of gene expression. We therefore used chromVAR[93] to assess whether
369 GDCA altered the accessibility of any TF motifs from the "CORE" collection of the well-established
370 JASPAR[94] database. We found that TF motif accessibility patterns quantified by chromVAR
371 clearly separated GDCA-exposed brains from controls (Fig. S11A). The TF motif with the highest
372 variability across samples was MA0007.4, which is recognized by the *Ar* transcription factor that
373 coordinates the cellular response to androgens (Fig. 5F; Supplementary Table 11). This *Ar* motif
374 was consistently less accessible in the GDCA-exposed brains than in controls (Fig. S11B).
375 A recent study showed that certain microbiome-derived bile acids, which share the sterol
376 structure of androgens, can antagonize the human androgen receptor[95]. Thus, GDCA may block
377 mouse *Ar* from nuclear translocation and binding to its target motifs, thereby reducing motif
378 accessibility - since TF binding usually opens chromatin[96] - and suppressing androgen-responsive
379 transcription. Consistent with this model, several of the genes strongly downregulated by GDCA
380 (e.g., *Adamts1*, *Akap12*, *Sgk1*) belong to the "Androgen Response" hallmark gene set from
381 MSigDB. GSEA likewise showed negative enrichment of this gene set (normalized enrichment
382 score = -2.17; Supplementary Table 8), although this did not reach statistical significance after

multiple testing correction (raw p = 0.05, adjusted p = 0.16), possibly reflecting a lag between certain chromatin changes and downstream transcriptional effects.

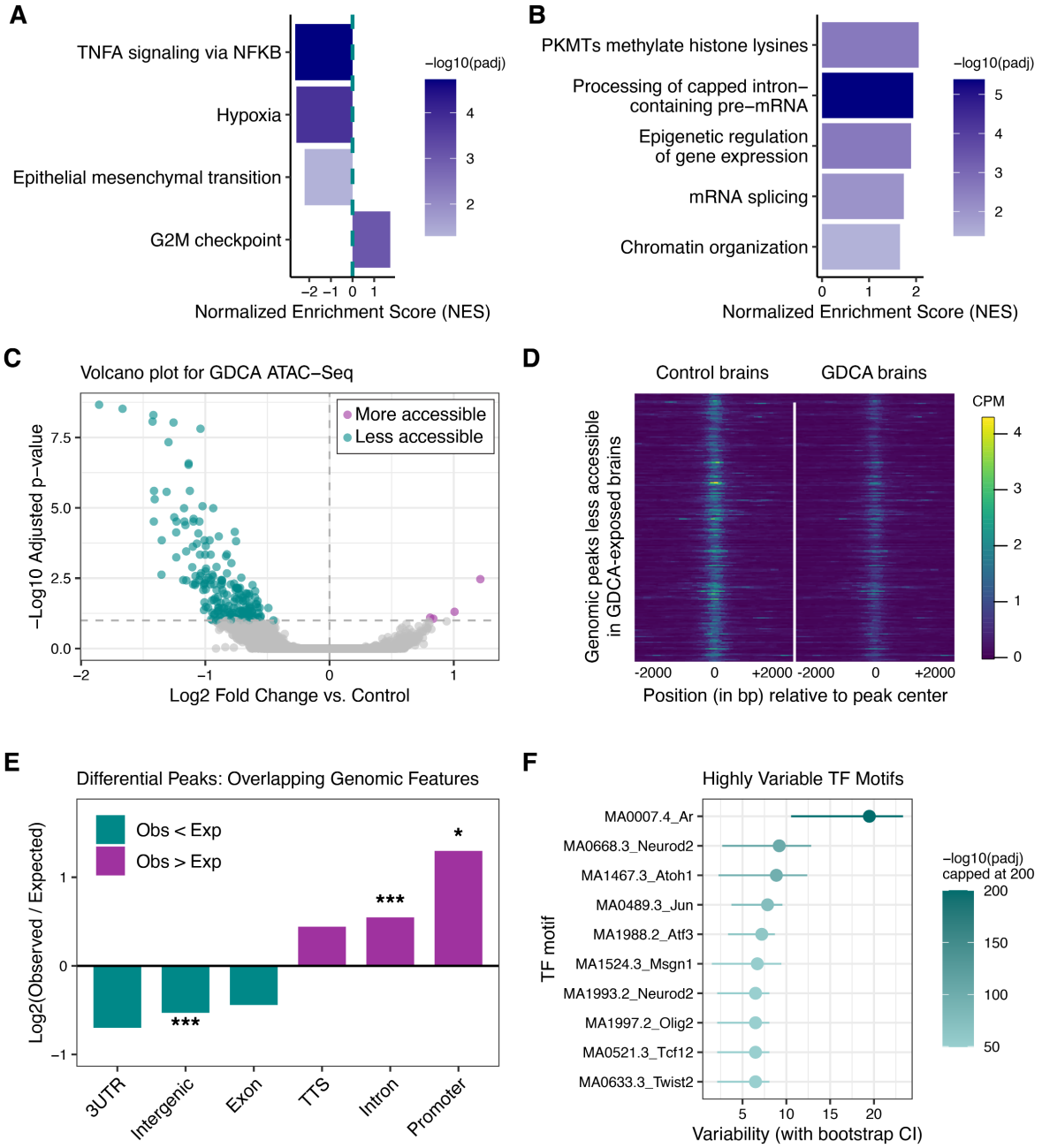


Fig. 5. GDCA suppresses stress-response genes and remodels chromatin in the brain. **(A)** All significantly enriched terms from gene set enrichment analysis (GSEA) based on hallmark gene sets; full results in Supplementary Table 8. **(B)** All significantly enriched terms from GSEA using the REACTOME pathway database; full results in Supplementary Table 8. **(C)** Volcano plot of differential chromatin accessibility between GDCA-treated and vehicle control mouse brains, based on ATAC-seq of flash-frozen right hemispheres from the initial RNA-seq cohort (see Methods). Y-axis: -log10(adjusted p-value) from DESeq2; x-axis: log2FoldChange vs. vehicle

394 control. Purple dots represent genomic peaks with higher accessibility in brains from GDCA-
395 exposed mice (adjusted $p \leq 0.1$ and positive log2FoldChange). Teal dots represent genomic
396 peaks with lower accessibility in brains from GDCA-exposed mice (adjusted $p \leq 0.1$ and negative
397 log2FoldChange). **(D)** Coverage heatmap for the 200 genomic peaks that were significantly less
398 accessible in brains from GDCA-exposed mice, vs. control. Each row represents one peak. BAM
399 files for individual brains within each condition were merged prior to visualization (Methods). CPM
400 = counts per million. **(E)** Observed versus expected proportions of differentially accessible peaks
401 overlapping various genomic features. Expected proportions were calculated based on the
402 reference genome. Introns and promoters are significantly enriched among differential peaks,
403 whereas intergenic regions are depleted. The symbol * indicates $p < 0.05$ and *** indicates $p <$
404 0.001. *p*-values were calculated by HOMER using a one-sided binomial test (upper-tailed when
405 observed > expected and lower-tailed when observed < expected). **(F)** Transcription factor (TF)
406 motifs with the highest variability in accessibility across ATAC-seq samples, as quantified by
407 chromVAR[93]. Dots represent point estimates of motif variability and horizontal lines indicate the
408 corresponding 95% bootstrap confidence intervals. The color scale shows the significance of
409 variability as -log10(adjusted *p*-value from chromVAR). To facilitate visualization, -log10(padj)
410 values were capped at 200; the true padj for the variability of MA0007.4_Ar across samples is 0.

411 GDCA acutely reduces mobility

412 Growing evidence indicates that androgens support both cognitive and motor function[97–99]. In
413 addition, increased chromatin accessibility can be beneficial for cognitive function[100], in stark
414 contrast to the accessibility reductions observed after GDCA treatment. Therefore, we
415 hypothesized that GDCA impairs cognitive and motor performance and may also influence
416 additional aspects of mouse behavior.

417 To test this, we first administered GDCA for three days (0.4 mg/kg/day, intravenous retro-
418 orbital injection) and assessed behavior (Fig. 6A). There were no detectable differences between
419 GDCA- and vehicle-treated mice, as assessed in an activity chamber on day 4 or in a Y-maze on
420 day 5 (n=15 male mice per group; see Figs. 6C-E, 6G-H, and Supplementary Table 12). We
421 reasoned that GDCA may have been cleared by the day of testing, and therefore evaluated the
422 impact of acute exposure in a second mouse cohort (Fig. 6B). 10 or 50 mg/kg GDCA was
423 administered via oral gavage, based on prior studies evaluating its impact on obesity and
424 polycystic ovary syndrome[101,102], and previous estimates that only a small percentage (5-10%) of
425 bile acids escape the enterohepatic circulation[103,104].

426 One hour after administration of GDCA or vehicle, mice were tested in the Y-maze. Like
427 vehicle-treated mice, GDCA-treated mice alternated between maze arms above chance levels
428 (Fig. 6F), indicating no impairment in spatial working memory. However, GDCA reduced mobility
429 in a dose-dependent manner, as evidenced by fewer entries into maze arms and lower total
430 distance moved within the Y-maze (Figs. 6I-L; Supplementary Table 12). At 50 mg/kg GDCA, this
431 was accompanied by a reduction in core body temperature (Fig. 6B and 6M). Total distance
432 moved also trended towards a GDCA-induced reduction within the setup of cohort 1 (p=0.054;
433 Fig. 6H; Supplementary Table 12). These results are intriguing given the association between
434 GDCA and Parkinson's Disease[32], which is characterized by both impaired locomotion[105] and a
435 reduced mesor (rhythm-adjusted mean) of core body temperature[106].

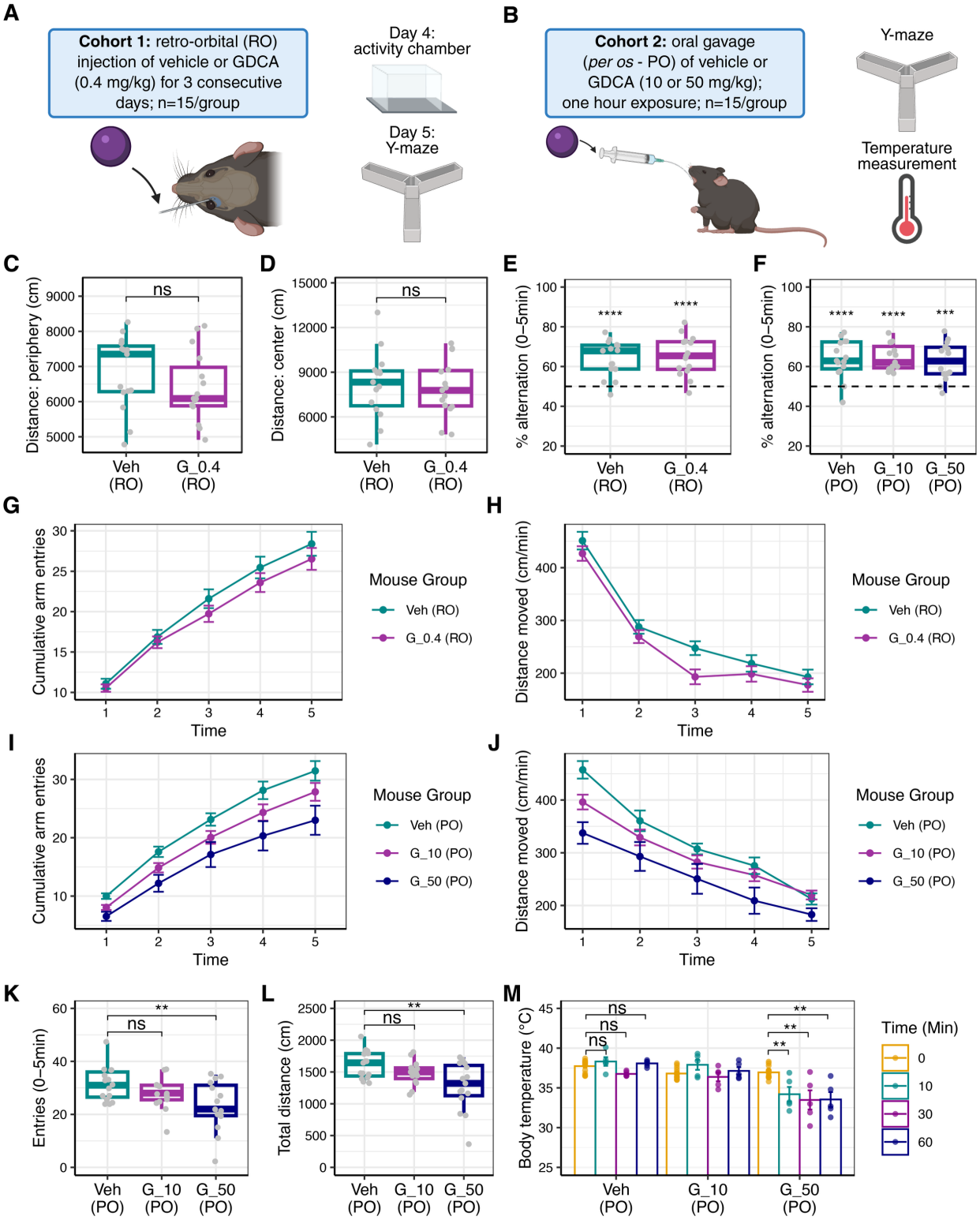


Fig. 6. GDCA acutely reduces mobility. **(A)** Schematic depicting the initial behavioral experiments performed. There were *n*=15 mice per group, yielding *n*=30 mice total in this first behavioral cohort. **(B)** Schematic depicting the second set of behavioral experiments. There were *n*=15 mice

per group, yielding *n*=45 mice total in this second behavioral cohort. Core body temperature was also measured in this cohort. **(C)** Total distance moved in the periphery of the activity chamber (cohort 1). Groups were compared using Welch's two-sample, two-sided *t*-test. **(D)** Same as (C), but for total distance moved in the center of the activity chamber. **(E)** Percent alternation between maze arms over the 5-minute duration of the Y-maze test (cohort 1). An alternation occurs when, after exiting one maze arm, the mouse does not enter one of the two arms they just visited, but rather enters the novel arm. Alternation percentages are defined as 100% * (number of alternations) / (total number of opportunities to alternate). Statistics compare the alternation percentages within each group to the chance level (50%[107,108]), using Welch's one-sample, two-sided *t*-test. **(F)** Same as (E), but for cohort 2. One mouse from the G_50 group was excluded as it had less than three arm entries and thus had no opportunities to alternate. **(G)** Cumulative entries into the arms of the Y-maze (cohort 1). **(H)** Distance moved within the Y-maze, minute-by-minute (cohort 1). **(I)** Same as (G), but for cohort 2. **(J)** Same as (H), but for cohort 2. **(K)** Boxplot of the data from minutes 0-5 in panel (I). Groups were compared using one-way ANOVA with Dunnett's post-hoc test. **(L)** Boxplot depicting the total distance traveled in the Y-maze by cohort 2, calculated as the sum of the values underlying panel (J). Statistics as in (K). **(M)** Core body temperature (°C) in cohort 2 measured at 0, 10, 30, and 60 min after gavage, six days after Y-maze testing. Bars show mean ± SEM. All *n*=15 mice per group were measured at time 0; *n*=5 mice per group were measured at each subsequent timepoint. Within each treatment group, time effects were assessed using one-way ANOVA; when significant, Dunnett's posthoc test was also applied (statistics shown). In all panels, ns indicates not statistically significant, * indicates $p \leq 0.05$, ** indicates $p \leq 0.01$, *** indicates $p \leq 0.001$, and **** indicates $p \leq 0.0001$. In all boxplots, boxes represent the interquartile range (IQR), the horizontal line indicates the median and whiskers extend between (25th percentile - 1.5*IQR) and (75th percentile + 1.5*IQR).

Discussion

The gut microbiome produces and impacts the systemic concentration of hundreds of metabolites, each of which may profoundly influence host organs, including the brain. However, systematic, high-throughput efforts to identify metabolites relevant to gut-brain signaling are lacking, partly because testing hundreds of molecules in animal models - the most accurate system to assess brain effects - is impractical. In this study, we integrated computational, *in vitro*, and *in vivo* approaches to identify age-related, microbiome-derived blood metabolites that - at physiologically relevant levels - induce significant molecular changes in the mouse brain. Among other changes, we observed impacts on RNA splicing and chromatin accessibility, whose dysregulation is implicated in brain aging and in age-related neurodegeneration[70,109,110]. Furthermore, we found that a microbiome-derived secondary bile acid, GDCA, can reduce mobility, a function that declines with aging and neurodegenerative disease. These findings demonstrate that "microbiome aging" may substantially contribute to "brain aging" through impacts on the blood metabolome, and that targeting the microbiome and its metabolites may offer cognitive benefits during aging.

The aminoxide TMAO has been extensively studied in the context of cardiovascular disease[13,111], and there is a growing body of research on its impact on the brain, where it has been shown to induce mitochondrial dysfunction, oxidative stress, neuroinflammation, and reductions in synaptic plasticity[29,58]. We found that TMAO also upregulates the glycolytic pathway in the brain - perhaps to compensate for mitochondrial dysfunction - and impacts RNA splicing by inducing strong downregulation of splicing factors and driving dozens of differential splicing events. Although prior work has identified splicing changes in the brains of germ-free vs. conventionally raised animals[112], to our knowledge, this is the first demonstration that specific microbiome-derived metabolites can alter host splicing patterns - whether in the brain or in any other host organ. GDCA may also affect splicing (Fig. 5B), raising questions about the relative impacts of each individual microbiome-derived blood metabolite and how these impacts might converge to influence normal brain function and the development of brain pathologies.

Unlike TMAO, little is known about the effects of GDCA on the mammalian nervous system, with prior knowledge limited to its association with neurodegenerative disease and cognitive impairment in several human cohorts[30–34]. Here, we build upon this knowledge, finding that GDCA alters chromatin accessibility, which might be the mechanism by which it also silences genes required for stress and injury response in the mouse brain. Large microbiome perturbations, such as antibiotics and germ-free status, have been shown to impact the brain epigenome[113]; however, previous metabolite-specific work has focused on the impact of short-chain fatty acids[114], leaving open questions about the many other classes of microbial effectors, including bile acids.

Our data suggest that GDCA substantially alters the brain transcriptome and epigenome, largely by antagonizing the host androgen receptor. This may suppress host androgen signaling, a pathway that is generally considered beneficial for cognitive and motor function[97–99]. This hypothesis aligns with our finding that GDCA exposure induces mobility deficits. However, further study - with varied dosing and testing paradigms - will be required to fully elucidate the cognitive and motor impacts of GDCA.

Although our work clearly demonstrates that circulating microbiome-derived metabolites can induce significant molecular alterations in the brain, it has several limitations. First, we tested only short-term exposures. However, brain aging and neurodegeneration occur over long periods of time. It will thus be important to decipher the short- vs. long-term effects of metabolites of interest. *In vitro*, this may require determining the half-life of each metabolite and ensuring continuous exposure to the desired concentration. *In vivo*, metabolites may need to be administered via osmotic minipump or through the diet. To determine the most physiologic dosing regimens, pharmacokinetic (PK) studies - involving metabolomics at various timepoints - will be necessary. Indeed, even the limited effects of *acute* exposure to PA and IAA in our study (Fig. 3B) may simply reflect suboptimal bioavailability of the administered formulations, underscoring the value of PK studies.

Another limitation is that, while our results show that TMAO increases the expression of glycolytic enzymes, both *in vitro* and *in vivo*, we did not definitively establish changes in glycolytic flux. This can be investigated using ^{13}C isotope tracing or similar approaches. Additionally, while we find interesting molecular phenotypes in adult mice, these results may or may not translate to aged animals. Future *in vivo* work in aged mice will be informative, in this regard, given that age may modify the brain's response to each metabolite. Relatedly, age-related comorbidities such as diabetes and cardiovascular disease may impact host response to individual metabolites, highlighting the importance of testing metabolite effects in the relevant disease models. Finally, because our multi-omic analyses were conducted using bulk rather than single-cell methods, the cell-type-specific effects of TMAO and GDCA remain unclear.

Despite these limitations, our work underscores that the gut microbiome, which is potentially modifiable, is a valuable therapeutic target in the context of brain aging. We identify specific age-associated microbiome metabolites that impact the transcriptional and epigenetic landscape of the brain, with likely consequences for brain function, as demonstrated by the observed impact of GDCA on locomotion. In addition, although we focused on brain aging in this work, the framework we describe - beginning with metabolomics data analysis, followed by *in vitro* transcriptomic screening, and concluding with *in vivo* studies of host organs and behavior - can be applied to identify the microbiome metabolites relevant to any host pathology of interest, providing a powerful tool for the field of microbe-host interactions.

Methods

Identifying microbiome-associated metabolites whose blood levels change with age

To identify metabolites whose blood levels change with aging, we leveraged data from three large cohort studies[18–20] listed in a recent review[17] on the metabolomics of aging. From Menni *et al.*[18], we retained all metabolites from Table S2, since only metabolites significantly associated with age were included in this table. From Dunn *et al.*[19], we retained metabolites with p adj <= 0.05 under the test labeled "_age.g2" in Supplementary Material 2 ("AGE" tab). From Darst *et al.*[20], we retained metabolites with p_age <= 0.05 in Table S1.

Separately, we curated a list of 329 human-associated metabolites that have been reported to be produced by microbes and/or are the products of host-microbe cometabolism. To create the list, we first downloaded Table S3 from a recent study on interactions between human-associated metabolites and GPCRs[21], and filtered to metabolites that included "Bacteria" in the "Source" column. After filtering, there were 320 metabolites. We then added the following nine metabolites known to be microbially producible and/or producible through host-microbe cometabolism[23,115–118]: acetic acid, butyric acid, indole-3-carboxaldehyde, indolelactic acid, enterolactone, imidazole propionate, p-cresol sulfate, hippuric acid and indoleacetic acid.

Then, we intersected the age-associated metabolites from each study with the curated list of metabolites associated with microbial metabolism. The intersection method was different per study given the differing formats of the tables. Specifically, for Darst *et al.*, we were able to intersect based on HMDB ID. We also noted that there was one metabolite that matched on CAS number but not on HMDB ID, given that it was listed as "ornithine HCl" in our curated list of microbiome-associated metabolites, but simply "ornithine" in the Darst study. So, we added ornithine manually to the intersection output. For the Menni and Dunn datasets, which lacked standardized identifiers, we used fuzzy string matching (via the Python package fuzzywuzzy, v0.18.0) to identify metabolites with similar names, followed by manual curation to confirm true matches. We also added two intersecting metabolites that were missed by fuzzy matching - indoleacetate and hippurate.

This process yielded 114 overlapping metabolites, 79 of which were unique based on HMDB ID. To narrow the list to metabolites most strongly associated with the gut microbiome, we leveraged two additional resources: **(1)** a recent study comparing blood metabolite levels between germ-free and conventionally raised mice[12]; and **(2)** the GUTSY Atlas[22], which reported the proportion of variance in blood metabolite levels "explained" by the gut microbiome for 1,168 metabolites based on paired stool-blood data for over 8,000 humans. Metabolites were retained for the final panel if they met at least one of two criteria: **(1)** showed significantly different blood levels between germ-free and conventionally raised mice[12]; **(2)** ranked in the top 20% of blood metabolites for the GUTSY variance-explained metric. We eliminated alanine and uridine, the former due to its minimal expected impact on cellular processes[24], and the latter because its blood levels increased with age in one study[18] but decreased with age in another[20].

The final table of 30 metabolites and their properties is contained in Supplementary Table 1. The table also specifies the exact product that was ordered for experimental testing.

582 Of note, at a later date, after conducting the *in vitro* RNA-seq experiments, we discovered
583 that the p_age values in Table S1 of the Darst study had not been corrected for multiple
584 hypothesis testing. However, even after we manually corrected for multiple hypothesis testing
585 (using a False Discovery Rate approach in R v4.3.2), all the metabolites from the Darst study in
586 our final panel maintained significance using an FDR threshold of padj ≤ 0.1, with all but cholate
587 (padj = 0.065) and glycodeoxycholate (padj = 0.061) also having a padj ≤ 0.05.

588 Cell culture

589 Human cerebral microvascular endothelial cells (hCMEC/D3; Sigma, Cat #SCC066) were
590 obtained as a gift from the lab of Prof. Mark Kay at Stanford University. Cell line authentication
591 was conducted using the ATCC Human Cell STR Profiling Service at passage #9 in October
592 2023, prior to initiating the experiments described in this manuscript. Cultures were maintained in
593 a 37°C, 5% CO_2 incubator at passage numbers ≤ 22, well below the limit of 35 established by the
594 developers of the cell line[26]. The cell medium was prepared from the EGM2-MV BulletKit (Lonza,
595 Catalog #CC-3202). The VEGF supplement was excluded from the medium based on literature
596 precedent[58,116] and due to VEGF's ability to impair the key barrier-forming properties of the
597 cells[119]. Cells were passaged every two days and were maintained on tissue culture-treated
598 plates. Cells were tested for mycoplasma contamination at six-month intervals during periods of
599 active culture using the ATCC Universal Mycoplasma Detection Kit (ATCC, Cat #30-1012K).

600 *In vitro* transcriptomic screen - experimental details

601 The screen was conducted in four separate batches using 12-well plates. Prior to cell seeding,
602 each well was coated with 0.5ml of a 1:20 dilution of rat tail collagen (Sigma, Catalog #08-115) in
603 DPBS (Fisher, Catalog #MT21031CV). Collagen was incubated on the plates at 37°C for 1 hour
604 and rinsed off twice with DPBS before cell seeding. Cells were seeded at passages #9-10, at a
605 density of 100,000 cells per well. Seeded cells were grown for six days prior to treatment, based
606 on prior literature[58,116]. The cell culture medium was refreshed every two days.
607 Each 12-well plate was organized to include three wells for vehicle controls and nine wells
608 for testing three metabolites, with three biological replicates per metabolite. All metabolites tested
609 on a given plate were dissolved in the same solvent (H_2O or DMSO) and vehicle controls included
610 an equivalent volume (1%) of the solvent in the final media used for the assays. Because urate
611 was insufficiently soluble in H_2O or DMSO alone, the final treatment media for this metabolite
612 included a small percentage (approximately 0.1%) of NaOH in addition to 1% H_2O.
613 Each batch consisted of three 12-well plates, which allowed room for testing nine
614 metabolites per batch. To allow for inclusion of an LPS control condition per batch, eight candidate
615 metabolites were tested per batch. Specifically, batches #2-4 tested 8 metabolites each and batch
616 #1 tested 6, to cover the entire set of 30 metabolites (see Supplementary Table 3). Cells were
617 treated with a metabolite, vehicle, or LPS for 3h before being harvested in TRIzol and stored at -
618 20°C for downstream RNA extraction and sequencing.
619 For each metabolite, we determined a physiologically relevant concentration to add to
620 cells (Supplementary Table 1), primarily based on information listed in the Human Metabolome
621 Database (HMDB) as of November 2023. Specifically, for each metabolite, we searched for the
622 metabolite in HMDB, clicked the "Concentrations" tab of the relevant entry, examined the

concentrations listed for “normal” adult blood samples, and selected a concentration based on those listings. HMDB did not provide blood concentrations for four metabolites: glycodeoxycholate, imidazole propionate, methyl indole-3-acetate, and N-acetyl-L-alanine. For glycodeoxycholate, we chose the concentration based on Table S2 of Qing *et al.*, *Schizophrenia* 2022[120]. For imidazole propionate, a relevant physiological concentration was determined based on Fig. 1 of Molinaro *et al.*, *Nature Communications* 2020[115]. For methyl indole-3-acetate, we were not able to find literature on the exact concentrations in adult blood samples. We tested 1μM because it was between previously reported concentrations of the metabolite in adult urine[121] and the chosen test concentration of the related metabolite indole-3-acetate. For N-acetyl-L-alanine, we tested 3μM based on previous literature that included *in vitro* experiments with this metabolite (Yang *et al.*, *Int J Mol Sci* 2023; Section 2.2[122]).

RNA was extracted using the following protocol. Samples were thawed to room temperature and incubated for an additional five minutes after thaw. 200μl chloroform was then added per 1ml of TRIzol used; then, samples were briefly vortexed and incubated 2-3 minutes at room temperature, prior to a 15-minute centrifugation at 17,000 x g at 4°C. The aqueous phase was transferred to a new tube. Steps 2-7 of the “Part 1” quick-start protocol for the RNeasy Mini Kit (Qiagen, Catalog #74106) were then executed, skipping the on-column DNase digestion but including the optional dry spin. DNase digestion was conducted using TURBO DNase (Fisher Scientific, Catalog #AM2239). Then, the “RNA cleanup” section of the RNeasy Mini handbook was executed, using 2 minutes rather than 1 minute for the dry spin. Finally, an ethanol precipitation was conducted for all samples to clean up any ethanol contamination, based on the relevant protocol from Cold Spring Harbor Laboratories[123]. RNA samples were sent to Novogene for quality control, mRNA library preparation (poly A enrichment), and sequencing (NovaSeq 6000, 150bp paired end reads, 6Gb raw data per sample).

In vitro transcriptomic screen - computational analysis

Computational analysis of the *in vitro* RNAseq data was performed separately for each sequencing batch, as follows. First, FASTQ files from Novogene were trimmed to remove adapter and low-quality bases using TrimGalore v0.6.10 with the following parameters: -q 30 --fastqc --stringency 3 --max_n 15 --output_dir trimming_out --cores 2 --retain_unpaired --paired <SAMPLE>_1.fq.gz <SAMPLE>_2.fq.gz.

Trimmed paired reads were pseudoaligned to the human reference transcriptome using kallisto v0.48.0[124]. The reference transcriptome was downloaded from Ensembl (release 110, Homo_sapiens.GRCh38.cdna.all.fa.gz) and an index was built using the kallisto index command. For pseudoalignment, the kallisto quant command was used with the -b 100 option, corresponding to 100 bootstraps.

Samples were then grouped based on the solvent included in the media (DMSO or H_2O), and a principal component analysis (PCA) was performed separately for each solvent group, as follows. First, transcript abundance tables generated by kallisto were normalized and filtered using the kallisto_table function from the sleuth R package (v0.30.2), using the options normalized = TRUE and use_filtered = TRUE. The resulting tables were log-transformed using the natural logarithm of (x+0.5) prior to PCA, which was conducted using the prcomp function in R v4.3.2.

PCA plots were used to identify potential outlier samples within each group - defined as samples positioned far from: (1) the vast majority of samples in the plot, and (2) the other

666 replicates within their treatment condition. Due to the subjective nature of visual outlier
667 identification, samples were only excluded from downstream analyses if their removal was
668 additionally supported by deviant quality control (QC) statistics from Novogene. **Note:** The PCA
669 plots for each sequencing batch are displayed in Figs. S4 and S5, and the corresponding sample
670 lists and QC statistics are contained in Supplementary Table 3.
671 Based on the above criteria, the following 5 (out of 138) samples were excluded from
672 downstream analysis. From batch 2, oxalate replicate #1 and carnitine replicate #2 were removed
673 (supporting QC statistic: low %GC content compared to the other samples in the batch). From
674 batch 3, glycodeoxycholate replicate #1 and quinolinate replicate #2 were removed (supporting
675 QC statistic: low %GC content compared to other samples in the batch). From batch 4 - DMSO
676 control replicate #5 was removed (supporting QC statistic: low RNA concentration compared to
677 all other DMSO control samples).
678 Following PCA and outlier removal, DESeq2[125] v1.42.1 was used to determine
679 differentially expressed genes between each set of metabolite-treated samples and the
680 corresponding vehicle controls in the sequencing batch, using the default significance cutoff of
681 alpha=0.10. As specified in Supplementary Table 3, the number of vehicle controls per solvent
682 per sequencing batch ranged from three to nine. Gene abundance values were determined from
683 the per-transcript kallisto outputs using tximport (v1.31.0). When applicable, the corrplot R
684 package (v0.95) was used to verify that all vehicle controls of the same solvent (e.g., 1% DMSO
685 controls) on different plates could be grouped, based on strong correlations - Pearson's $r > 0.98$
686 - of log-transformed gene abundances.
687 As recommended by the DESeq2 developers, the function lfcShrink[60] was applied to
688 conservatively shrink log2foldchanges, using the shrinkage estimator 'apeglm.' Additionally, to
689 avoid overcounting the number of significant hits per metabolite, we filtered the DESeq2 output of
690 differentially expressed genes to eliminate lines that were exact duplicates of previous lines
691 besides the Ensembl ID field, which can occur when two different Ensembl gene IDs are mapped
692 to the same physical gene. Finally, genes with an absolute shrunken log2 fold change ≤ 0.1 were
693 not counted as DEGs, even if they met the significance threshold of adjusted $p \leq 0.1$.
694 The KEGG pathway enrichment analyses shown in Figs. 2B and S2 were conducted using
695 g:Profiler (https://biit.cs.ut.ee/gprofiler/gost)[126] on August 17, 2025 using the default g:SCS
696 significance threshold and a custom list of background genes - specifically, all genes that passed
697 DESeq2's internal filters in the corresponding DESeq2 analysis. For the analysis of DEGs in the
698 Alzheimer and Parkinson's KEGG pathways (Fig. 2C), the pathway gene lists were retrieved on
699 July 19, 2025 from https://rest.kegg.jp/link/hsa/path:hsa05010 and
700 https://rest.kegg.jp/link/hsa/path:hsa05012, respectively. The numbers reported in Fig. 2C
701 represent numbers of unique genes (based on Entrez ID).

702 *In vitro* cytotoxicity testing

703 hCMEC/D3 cells (passage #22) were seeded on Day 0 at a density of 100,000 cells per well in
704 12-well plates. Prior to seeding, each well was coated with 23µg rat tail collagen (Sigma, Cat.
705 #08-115) in 1mL DPBS (Fisher, Catalog #MT21031CV) and rinsed twice with DPBS following a
706 10-minute incubation at room temperature.
707 On Day 2, cells were treated for 3 hours with EGM2-MV BulletKit medium (Lonza, Catalog
708 #CC-3202) excluding the VEGF supplement and supplemented with either: **[1]** metabolites at the

709 concentrations shown in Fig. S7, **[2]** vehicle control (i.e., 1% DMSO or 1% H_2O), or **[3]** 20%
710 DMSO (cytotoxic control).
711 Following treatment, cells were detached with 0.05% trypsin-EDTA (Fisher Scientific,
712 Catalog #25-300-054), pelleted by centrifugation (100 x g, 5min, 4°C), and resuspended in 150µl
713 serum-free, phenol red-free EGM2-MV medium (basal medium: Lonza, Catalog #CC-3129;
714 supplements: Lonza EGM2-MV SingleQuots kit, Catalog #CC-4147, excluding FBS and VEGF).
715 30µl of Propidium Iodide ReadyFlow reagent (Thermo Fisher, Catalog #R37169) was added to
716 the suspension, and fluorescence was measured using a NovoCyte Quanteon flow cytometer at
717 the Stanford Shared FACS Facility.

718 Animals

719 Male C57BL/6 mice (3 months old) were obtained from Jackson Laboratories (stock # Jax
720 000664). All mice were kept on a 12 h-12 h light-dark cycle and provided *ad libitum* access to food
721 and water. All animal care and procedures complied with the Animal Welfare Act and were in
722 accordance with institutional guidelines and approved by the institutional administrative panel of
723 laboratory animal care at Stanford University (PIs of animal protocols: T. Wyss Coray, M.
724 Shamloo; protocol #s 29100 and 18466).

725 Mouse brain transcriptomics after metabolite injection

726 Guided by prior studies investigating TMAO and the microbiome metabolite p-Cresol sulfate[58,116],
727 we conducted bulk transcriptomics on mouse half-brains following a 2-hour exposure to each
728 metabolite of interest or the corresponding vehicle control. As detailed in Figs. #2-3 and the
729 accompanying text, four metabolites were tested in mice - palmitic acid (PA), indoleacetic acid
730 (IAA), trimethylamine N-oxide (TMAO), and glycodeoxycholic acid (GDCA). Intravenous retro-
731 orbital injection under isoflurane anesthesia was used as the administration route.
732 Metabolite dosing was based on precedent from the study on TMAO that guided
733 experimental design[58]. For TMAO, we used the same dose as the original study, 1.8 mg/kg. This
734 dose was described as "calculated to approximate human circulating TMAO levels." Thus, for the
735 remainder of the metabolites, we examined the targeted metabolomics literature[32,34,48–55] to
736 identify the range of circulating levels in human seniors, and calculated mouse-equivalent doses
737 of observed human concentrations, using well-established guidelines[56,57]. PA (Sigma, Catalog
738 #P0500-10G, Lot #0000253065) was administered at a final dose of 22.2 mg/kg (~100µM total
739 body dose). IAA (Sigma, Catalog #I2886-5G, Lot #0000249926) was administered at a final dose
740 of 1.5 mg/kg (~10µM total body dose). GDCA (Sigma, Catalog #361311-5GM, Lot #4045885) was
741 administered at a final dose of 0.4 mg/kg (~1µM total body dose).
742 Two experiments were conducted in total. **Experiment 1** contained 8 mice; 4 mice
743 received 1.8 mg/kg TMAO in 100µl saline, and 4 mice received 100µl saline (vehicle control).
744 **Experiment 2** contained 20 mice, in two groups. Group 1 contained 8 mice; 4 mice received 0.4
745 mg/kg GDCA in 100µl saline, and 4 mice received 100µl saline (vehicle control). Group 2
746 contained 12 mice; 4 mice received 22.2 mg/kg PA in 100µl saline containing 6% DMSO, 4 mice
747 received 1.5 mg/kg IAA in 100µl saline containing 6% DMSO, and 4 mice received 100µl saline
748 also containing 6% DMSO (vehicle control). ***Note:*** unlike TMAO and GDCA, PA and IAA are only
749 soluble in organic solvents such as DMSO, hence the use of DMSO above.

750 For sacrifice, mice were anesthetized using 2.5% Avertin (v/v). As per the original TMAO
751 study[58], mice were transcardially perfused with 0.9% saline after sacrifice and prior to brain
752 removal, in order to remove contaminating blood. Forceps used for brain extraction were pre-
753 treated with RNaseZap (Thermo Fisher, Catalog #AM9780) and subsequently rinsed twice with
754 RNase-free water to minimize the risk of RNA degradation. Brains were sagittally hemisected at
755 the corpus callosum prior to storage. Left hemispheres were stored in 1ml RNAlater overnight at
756 4°C, and then dabbed on a Kimwipe to remove excess RNAlater before transfer to a fresh tube
757 for storage at -80°C. Right hemispheres were either processed in the same manner or flash-
758 frozen directly at -80°C. Specifically, for the TMAO-treated and corresponding vehicle control
759 brains - which were part of the first experiment performed - the right hemispheres were processed
760 identically to the left. In the second experiment, however, the right hemispheres were flash-frozen
761 at -80°C without RNAlater incubation.
762 RNA was extracted from frozen left hemispheres using a modified version of the protocol
763 from the QIAWave RNA Mini Kit (Qiagen, Catalog #74536), as follows. A bead beater (Qiagen
764 Tissue Lyser II) was used to homogenize the left hemisphere in a collection microtube (Qiagen
765 Catalog #19560) containing 700µL of pre-chilled QIAzol (Qiagen, Catalog #79306) and stainless
766 5mm steel bead (Qiagen Catalog #69989). For bead beating, the tube was positioned in between
767 cold metal blocks prechilled at -20°C. Samples were homogenized at 30Hz for five minutes, kept
768 on ice for 1 minute, then homogenized at 30Hz for an additional 2 minutes. Following
769 homogenization, the tube was incubated at room temperature for 3-5 minutes. 200µL was then
770 set aside for RNA extraction in a DNA LoBind tube; the remaining volume was stored in a fresh
771 DNA LoBind tube at -80°C. 500µL of additional QIAzol was added to the 200µL tube to make
772 700µL; then, all the liquid was transferred to a DNA LoBind tube containing 200µL chloroform.
773 The tube was capped and vortexed for 15 seconds, then incubated at room temperature for 2-3
774 minutes and centrifuged at 12,000xg for 15min at 4°C. 150µL x 2 times (300µL total) of aqueous
775 phase was then transferred to a DNA LoBind tube containing 300µL 70% ethanol. The tube was
776 then inverted 10 times to mix, and briefly centrifuged. ~600µl of sample was transferred to the
777 RNeasy Mini spin column, then spun at 10,000xg for 30 seconds at room temperature. Flow-
778 through was discarded, and 350µL RW1 was added to the column. The column was spun at
779 10,000xg for 30 seconds at room temperature, and the flow-through was again discarded. 500µL
780 RPE was added to the column, and the column was spun at 10,000xg for 30 seconds at room
781 temperature, after which flow-through was discarded. The preceding RPE wash and spin was
782 then repeated for 2 minutes. The tube was then centrifuged at maximum speed for one minute at
783 room temperature. The column was transferred to a new 1.5ml DNA LoBind tube, to which 50µL
784 RNAse-free water was added. The lid was closed and the tube was incubated at room
785 temperature for 1-2 minutes. For elution, the tube was spun for one minute at 10,000xg at room
786 temperature. The same 50µL of eluate was used to repeat the previous step.
787 RNA samples were sent to Novogene for quality control, mRNA library preparation (poly
788 A enrichment), and sequencing (NovaSeq 6000, 150bp paired end reads, 6Gb raw data per
789 sample). A sample list and QC statistics are included in Supplementary Table 3.

790 Mouse brain transcriptomics - preliminary data analysis

791 The following steps were performed separately for each experiment defined above. First, FASTQ
792 files from Novogene were trimmed to remove adapter and low-quality bases using TrimGalore

v0.6.10 with the following parameters: --fastqc --stringency 3 --max_n 15 --output_dir trimming_out --cores 2 --retain_unpaired --2colour 30 --paired <SAMPLE>_1.fq.gz <SAMPLE>_2.fq.gz.

Trimmed paired-end reads were pseudoaligned to the standard mouse transcriptome index provided with the v1 release of kallisto[124], using kallisto v0.50.1. For pseudoalignment, the kallisto quant command was used with the -b 100 option, corresponding to 100 bootstraps.

Principal component analysis (PCA) was conducted using the following functions available in DESeq2 (v1.42.1): vst (setting blind = TRUE) and plotPCA. PCA revealed two key insights that guided subsequent decisions: (1) For experiment 1 - TMAO replicate #3 was positioned away from every other TMAO-treated sample along the first two components (Fig. S6A), necessitating further investigation of this sample prior to downstream analysis. (2) For experiment 2 - PCA revealed strong separation of samples by RNA extraction batch (Fig. S6B), necessitating the inclusion of this variable as a covariate in downstream differential expression analysis.

TMAO replicate #3 was ultimately excluded from subsequent analyses, including DESeq2 differential expression analysis, based on three additional facts beyond its placement in PCA. First, TMAO replicate #3 exhibited high leverage in the initial DESeq2 run; i.e., with TMAO replicate #3 included, it showed the highest mean Cook's distance, more than 1.5x that of any other sample, indicating a disproportionate influence on the linear model. Second, DESeq2 differential gene expression analysis comparing TMAO replicate #3 vs. all other samples revealed enrichment of blood markers (Hba-a1: $\log_2$FC = 2.28, padj = 3.9 × 10E-4; Hbb-bt: $\log_2$FC = 1.72, padj = 0.03; Supplementary Table 7), suggesting potential incomplete perfusion and/or other blood contamination. Third, excluding TMAO replicate #3 brought our differential expression results into closer agreement with those reported by Hoyles *et al.*[58], the study that inspired the design of this experiment, and with the broader literature on TMAO, strengthening confidence in the results.

DESeq2[125] v1.42.1 was used to determine differentially expressed genes between metabolite-treated samples and matched vehicle controls, using the default significance cutoff of alpha=0.10. For experiment 1, TMAO replicate #3 was excluded for the reasons detailed above, yielding *n*=3 TMAO brains and *n*=4 control brains for the final differential expression analysis. For experiment 2, all *n*=4 brains per condition were included in the final analysis, and the RNA extraction batch was included as a covariate, based on the PCA results described above (Fig. S6B). As recommended by the DESeq2 developers, the function lfcShrink[60] was applied to conservatively shrink log2foldchanges, using the shrinkage estimator 'apeglm.' All final DESeq2 differential gene expression results for PA, IAA, TMAO and GDCA are contained in Supplementary Table 7. Genes with an absolute shrunken log2foldchange ≤ 0.1 were not counted as DEGs, even if they met the significance threshold. All per-gene abundances for DESeq2 analysis were determined from the per-transcript kallisto outputs using tximport (v1.30.0).

Mouse brain transcriptomics - gene set enrichment analysis

Gene set enrichment analysis (GSEA) was conducted using the R package fgsea (v1.28.0). The relevant gene sets (hallmark gene sets and REACTOME pathways for mouse) were downloaded from https://www.gsea-msigdb.org/gsea/msigdb/mouse/collections.jsp on July 4, 2025. Enrichments with padj ≤ 0.05 were considered statistically significant.

Mouse brain transcriptomics - marker gene analysis

Marker genes for six different brain cell classes were determined using the following command in the R package Seurat (v5.3.0): FindAllMarkers(seurat_object, group.by = "ident", slot = “data”, only.pos = TRUE, min.pct = 0.25, logfc.threshold = 0.5). The resulting marker genes were further filtered to those with adjusted *p*-value ≤ 0.05. The Seurat object was constructed using single cell data from the mouse brain, specifically data from the 8 young mice (age-matched to those in our study) included in Ximerakis *et al.*, 2019[65]. Expression data, and metadata specifying the class of each cell, were downloaded from https://singlecell.broadinstitute.org/single_cell/study/SCP263/aging-mouse-brain on August 22, 2025.

Mouse brain transcriptomics - differential splicing analysis

Differential splicing events were identified using rMATS-turbo[71] (v4.1.1), as follows. First, STAR (v2.7.11b) was used to align trimmed paired-end reads to the mouse genome downloaded from GENCODE (Release M37, i.e., GRCm39). The following command was used: STAR --genomeDir <directory_containing_genome> --readFilesIn <SAMPLE>_1.trimmed.fq.gz <SAMPLE>_2.trimmed.fq.gz --outFileNamePrefix <SAMPLE> --outSAMunmapped Within --outSAMattributes NH HI AS NM MD XS --twopassMode Basic --alignSJDBoverhangMin 1 --alignSJoverhangMin 8 --runThreadN 16 --outSAMtype BAM SortedByCoordinate --outSAMstrandField intronMotif --readFilesCommand zcat.

The resulting bam files were then run through rMATS-turbo using the following command: rmats.py --gtf <path_to_genome_annotation_file> --b1 <file_specifying_saline_bams> --b2 <file_specifying_tmao_bams> --od <output_directory> --tmp <directory_for_tmp_files> -t paired --libType fr-unstranded --readLength 150 --variable-read-length --nthread 10 --cstat 0.1 --novelSS --allow-clipping. Based on the “JCEC” output files, this command initially nominated 126 differential splicing events between TMAO-treated and saline-treated brains (FDR ≤ 0.05), which we further narrowed to 44 using the following four filters: **(1)** ≥ 5 read counts on average per group, **(2)** exclusion of events with average PSI < 0.05 or > 0.95 in both groups, **(3)** FDR ≤ 0.01, and **(4)** ΔPSI ≥ 0.10.

The sashimi plot in Fig. S9A was generated using ggsashimi[127] (v1.1.5).

Bulk ATAC-seq of brains of GDCA-exposed vs. vehicle control mice

As described above under “Mouse brain transcriptomics after metabolite injection,” only the left hemisphere of each brain in these experiments underwent RNA-sequencing. The right hemispheres from GDCA-exposed and corresponding vehicle control mice were flash-frozen at -80°C, enabling ATAC-seq.

Mouse brain nuclei were isolated following procedures established for processing mammalian brains[128] and then ATAC-seq was carried out as previously described[129] with some modifications.

Briefly, dissected brain tissue was added to 1 mL nuclei extraction buffer (NEB: 250 mM Sucrose, 65 mM B-glycerol, 1x protein inhibitor, 25 mM KCl, 5 mM MgCl2, 20mM HEPES-KOH, 0.5% IGEPAL CA-630, 1 mM DTT, 0.2 mM Spermine, 0.5 mM Spermidine, 2-5% NGSerum) and

875 incubated on ice for 10-15 minutes. then dounced 20 times with pestle A and another 20 times
876 with pestle B. Another 300 µL of NEB was added and the mixture was transferred to a 2mL tube
877 and incubated on a rotator at 4°C for 5 minutes, then filtered through a 70µm Flowmi strainer
878 (Sigma, Catalog #BAH136800070) into a fresh 2mL tube. The volume was adjusted to 1,200 µL
879 with more NEB if necessary and nuclei were lightly crosslinked by adding 6.52 µl of 37% FA,
880 incubating for 4 minutes at room temperature, and quenching with 133 µL 2.5 M Glycine solution.
881 Nuclei were finally isolated using density gradient separation, by adding 1,100 µL of O-
882 50% solution (generated by mixing 18,751 µL O-60%/OptiPrep/iodaxinol solution, 3242 µL H20,
883 450 µL NGS, and 22.5 µL 1M DTT) in a 5mL Eppendorf tube, then creating a density gradient by
884 adding 1,000 µL O-44% (8,800 µL O-50% plus 1,200 µL H_2O) to the very bottom of the tube and
885 500 µL O-22% (2,200 µL O-50% plus 2,800 NEB buffer) in between the O-44% and the top layer
886 with the nuclei. Tubes were then centrifuged for 30 minutes at 4°C with breaks off (deceleration
887 0, acceleration 5), and nuclei were extracted from the middle layer in the density gradient.
888 After nuclei isolation, nuclei were counted and approximately 100,000 were used as input
889 into each transposition reaction in four technical replicates for each individual mouse brain.
890 Transposition was carried out in 50 µL transposition mix, consisting of 25 µL 2x TD buffer (20 mM
891 Tris-HCl pH 7.6, 10 mM MgCl2, 20% Dimethyl Formamide), 22.5 µL nuclease-free H_2O, and 2.5
892 µL loaded Tn5 transposase (produced and assembled following previously established
893 protocols[130] by incubation in an Eppendorf ThermoMixer at 37°C at 1,000 rpm for 30 minutes. The
894 transposition reactions were then stopped immediately by adding 150 µL reverse crosslinking
895 buffer (1% SDS, 0.1M NaHCO3), and reverse crosslinking was carried out for 8 hours at 65°C.
896 Transposed DNA was isolated using the MinElute PCR Purification Kit (Qiagen, Catalog
897 #28004) by adding 600 µL PB Buffer and eluting in 20 µL EB Buffer heated to 65°C, then PCR
898 amplified by mixing 20 µL transposed DNA, 2.5 µL i5 and 2.5 µL i7 PCR primers, and 25 µL 2x
899 NEBNext High-Fidelity PCR Master Mix (New England Biolabs, Catalog #M0541L), with the
900 following settings: 72°C for 5 minutes, 98°C for 30 seconds, followed by 10 cycles of 98°C for 10
901 seconds, 63°C for 30 seconds, and 72°C for 45 seconds.
902 Final libraries were purified using the MinElute PCR Purification Kit (Qiagen, Catalog
903 #28004), QC-ed using a Qubit 1X dsDNA kit (Thermo Fisher Scientific, Catalog #Q33230) and an
904 Agilent TapeStation (High Sensitivity D1000 Screen Tape® Assay), and sequenced on a NextSeq
905 550 instrument in a 2x38 format.

906 ATAC-seq - data analysis

907 ATAC-seq data analysis was conducted largely as previously described[131], with brief specifics
908 provided below.
909 Two replicates were sequenced per brain sample, yielding *n*=16 total replicates from *n*=8
910 mice, of which *n*=4 were GDCA-exposed mice and *n*=4 were the corresponding vehicle control
911 mice (Supplementary Table 3).
912 Forward and reverse reads from each replicate were first trimmed to a length of 36bp
913 using the script PEFastqToTabDelimited.py available at the following Github repository:
914 https://github.com/georgimarinov/GeorgiScripts. Trimmed reads were aligned to both the entire
915 mouse genome (mm10) and to the mitochondrial portion alone (chrM) using bowtie (v1.0.1). The
916 following parameters were used for alignment to the mitochondrial genome: -p 20 -v 2 -a -t --best

--strata -q -X 1000 --sam --12. The following parameters were used for alignment to the entire genome: -p 20 -v 2 -k 2 -m 1 -t --best --strata -q -X 1000 --sam --12.

The mitochondrial alignments were used to calculate the fraction of mitochondrial reads in each replicate (reported in Supplementary Table 3). For all subsequent analyses, the entire-genome alignments were used, after filtering out reads mapping to chrM.

The remaining, non-mitochondrial reads were deduplicated using the MarkDuplicates tool from Picard Tools (v1.99) using the following parameters: VALIDATION_STRINGENCY=LENIENT ASSUME_SORTED=true REMOVE_DUPLICATES=true. The resulting deduplicated .bam files were indexed using samtools v0.1.18 and converted to .wig using the script makewigglefromBAM-NH.py from the Github repository linked above. Reads per million (RPM) was used as the normalization method. The .wig files were converted to .bigWig using wigToBigWig from UCSC Utils (version from 2017-07-13). TSS enrichment scores reported in Supplementary Table 3 were calculated using the scripts signalAroundCoordinate-BW.py and ATACTSSscore.py, from the Github repository linked above. TSS coordinates were based on the UCSC mm10 refFlat gene annotation table (dated 2014-10-26).

Peaks were called for each replicate using the callpeak function from MACS2 (v2.1.1) using the following parameters: -g mm -f BAMPE --to-large -p 1e-1 --keep-dup all --nomodel. Then, reproducible peaks between replicates were identified using the previously described irreproducible discovery rate (IDR) framework[131]. IDR v2.0.4.2 was used for this step. The required script BAMPseudoReps.py from the Github repository linked above was converted to Python 3 using the 2to3 Python program prior to use.

The reproducible peaks between replicates per sample were then merged across samples using the createIterativeOverlapPeakSet.R script available at the following Github repository - https://github.com/corceslab/ATAC_IterativeOverlapPeakMerging - described in Grandi *et al.*, *Nature Protocols* 2022[132]. Blacklist regions provided to the script were downloaded from https://github.com/Boyle-Lab/Blacklist/tree/master/lists (file mm10-blacklist.v2.bed.gz) on July 11, 2025. The spm was set to 0 because confident peaks had already been identified. Additional relevant parameters were: --suffix _summits.bed --blacklist mm10-blacklist.v2.bed --genome mm10 --rule "(n+1)/2" --extend 250. This step yielded a .bed file with the coordinates of the final set of merged peaks, which was converted to SAF format and used to create a read counts table (rows = merged peaks, columns = samples) using featureCounts (v2.0.6) as follows: featureCounts -p --countReadPairs -B -C -T 8 -F SAF -a All_samples_peaks.saf -o All_Samples.counts <Brain1.pooled_replicates.sorted.bam> <Brain2.pooled_replicates.sorted.bam> <Brain3.pooled_replicates.sorted.bam>…

Samples were retained for subsequent analysis if both replicates had TSS enrichment scores ≥ 10, the standard deemed as “acceptable” under ENCODE guidelines for mm10 (https://www.encodeproject.org/atac-seq/, last accessed July 28, 2025). As per Supplementary Table 3, 7 samples (*n*=4 from GDCA-exposed mice, and *n*=3 from control mice) passed this filter.

Differentially accessible peaks were determined using DESeq2 (v1.42.1) using the default significance cutoff of alpha=0.10. These peaks were further annotated using the annotatePeaks.pl script from HOMER, which was installed using the configureHomer.pl script downloaded from http://homer.ucsd.edu/homer/ on July 14, 2025. Exact command used: annotatePeaks.pl

960 sig_peak_file_for_homer.txt mm10 -annStats annotation_stats.txt > annotated_sig_peak_file.txt.
961 The strand in sig_peak_file_for_homer.txt was specified as "." for all peaks.
962 Coverage of differentially accessible peaks (Fig. 5D; Fig. S10A-B) was determined as
963 follows. First, BAM files for individual brains within each condition were merged using samtools
964 v1.21. The resulting two files (one for control brains and one for GDCA-exposed brains) were
965 sorted and indexed with samtools v1.21, and subsequently converted to bigWig format using the
966 bamCoverage function from deepTools v3.3.0, using a bin size of 100 bp and specifying counts
967 per million (CPM) as the normalization method. Coverage matrices per condition were generated
968 from each bigWig file using the signalAroundCoordinate-individual.py script available from
969 https://github.com/georgimarinov/GeorgiScripts, using the previously specified parameters[131],
970 after conversion to Python 3 using the 2to3 Python program.
971 chromVAR[93] analysis (Figs. 5F and S11) was conducted using chromVAR 1.24.0,
972 referencing the 0.99.6 version of the JASPAR 2024 database in R version 4.3.2.

973 Behavioral testing and measurement of core body temperature

974 Behavioral testing was conducted at the Stanford Behavioral and Functional Neuroscience
975 Laboratory. Mice were acclimated to a reverse 12h-12h light-dark cycle for one week prior to
976 initiation of experiments, ensuring that daytime testing coincided with their naturally active (dark)
977 phase.
978 For experiments with mouse cohort 1 (schematized in Fig. 6A), mice were dosed for three
979 consecutive days with GDCA (0.4 mg/kg) or saline vehicle (n=15 mice per group, n=30 total).
980 Intravenous retroorbital injection was used as the administration route, with isoflurane as the
981 anesthetic. The total injection volume was 50μL. Behavioral testing was conducted after the final
982 injection (activity chamber: one day after final injection, Y-maze: two days after final injection).
983 Mice were observed for one hour in the activity chamber and for five minutes in the Y-maze.
984 For experiments with mouse cohort 2 (schematized in Fig. 6B), mice were administered
985 50 mg/kg GDCA, 10 mg/kg GDCA, or saline vehicle *per os* (n=15 mice per group, n=45 total). Y-
986 maze testing was conducted one hour post-dosing. Six days after Y-maze testing, all mice were
987 redosed (keeping the same group assignments), and core body temperature was measured using
988 a mouse rectal probe. All n=15 mice per group were measured at timepoint 0, prior to dosing; n=5
989 mice per group were measured at each subsequent timepoint (10min, 30min, and 60min post-
990 dosing).

991 # Data availability

992 All raw sequencing data are available under NCBI BioProject ID PRJNA1328687[133].

993 # Code availability

994 All software, parameters, and relevant details of custom code are described in the Methods
995 section.

Acknowledgements

We thank Jeff Rubin, Ren Song, Ayan Mondal, Hongchao Guo, Michael Bassik, Betsy Mellins, and Mark Kay for guidance surrounding brain endothelial cell culture and experiments. Jeff Rubin and Mark Kay provided us with the hCMEC/D3 cell line. We thank Jakub Rajnuk for advice during the metabolite selection process, Anshul Kundaje for guidance on RNA-seq data analysis, Garrett Patrick for reviewing the manuscript during its preparation, Ariel Gulasch for assistance with *in vitro* transcriptomic screening, and Rudy Wycallis at the Stanford Shared FACS Facility for assistance with flow cytometry. Several non-data figures such as experiment timeline figures were created using BioRender.com. This work was funded, in part, by NIH R01AI148623 and R01AI143757, NCI R01 CA301727, a Stand Up 2 Cancer Grant, and the Stanford Wu Tsai Neurosciences Institute. M.C. was supported by the National Defense Science and Engineering Graduate Fellowship. I.E.P. was supported by the Stanford REACH PhD fellowship. S.M.S. was supported by a Stanford Bio-X Bowes Fellowship. Computing costs were supported, in part, by an NIH S10 Shared Instrumentation grant (1S10OD02014101). The Stanford Behavioral and Functional Neuroscience Laboratory is supported by the NIH S10 Shared Instrumentation for Animal Research (1S10OD030452-01).

Author contributions

MC: conceptualization, investigation, methodology, data curation, formal analysis, visualization, writing - original draft, writing - review and editing. **SMS:** investigation, methodology. **IEP:** methodology. **DJR:** investigation, methodology. **GKM:** investigation, methodology, data curation, formal analysis. **AAM:** investigation, writing - review and editing. **JLEB:** investigation, methodology. **AN:** visualization. **JWJ:** methodology. **JCW:** supervision. **SXL:** methodology, supervision. **SMD:** methodology, supervision. **WJG:** supervision. **NLS:** investigation, methodology, data curation, formal analysis. **MS:** methodology, supervision. **AB:** funding acquisition, supervision. **TWC:** supervision. **ASB:** conceptualization, methodology, funding acquisition, supervision, writing - original draft, writing - review and editing. All authors read and approved the manuscript.

Competing interests

The authors declare no competing interests.

Supplementary information

Supplementary Information is available for this paper, as follows:

1. PDF of Supplementary Figures
2. Supplementary Table 1 - Properties of the 30 metabolites in the final screening panel
3. Supplementary Table 2 - Enriched KEGG pathways for genes upregulated by LPS

4. Supplementary Table 3 - Sample information for omics assays
5. Supplementary Table 4 - DESeq2 results, screen batches 1 and 2
6. Supplementary Table 5 - DESeq2 results, screen batches 3 and 4
7. Supplementary Table 6 - Differential genes in AD or PD pathways for 12 metabolites
8. Supplementary Table 7 - DESeq2 results, mouse brain RNA-seq
9. Supplementary Table 8 - GSEA results, mouse brain RNA-seq
10. Supplementary Table 9 - Significant differential splicing events in TMAO brains
11. Supplementary Table 10 - Studies associating GDCA in blood with AD, PD, or CI
12. Supplementary Table 11 - ATAC-seq data analysis results
13. Supplementary Table 12 - Full stats for behavioral and temperature data

References

1. Sachdev, P. S. *et al.* Factors predicting reversion from mild cognitive impairment to normal cognitive functioning: a population-based study. *PLoS One* **8**, e59649 (2013).
2. Shimada, H., Doi, T., Lee, S. & Makizako, H. Reversible predictors of reversion from mild cognitive impairment to normal cognition: a 4-year longitudinal study. *Alzheimers Res Ther* **11**, 24 (2019).
3. Badal, V. D. *et al.* The Gut Microbiome, Aging, and Longevity: A Systematic Review. *Nutrients* **12**, (2020).
4. Pang, S. *et al.* Longevity of centenarians is reflected by the gut microbiome with youth-associated signatures. *Nat Aging* **3**, 436–449 (2023).
5. Boehme, M., Guzzetta, K. E., Wasén, C. & Cox, L. M. The gut microbiota is an emerging target for improving brain health during ageing. *Gut Microbiome (Camb)* **4**, (2023).
6. Cryan, J. F. *et al.* The Microbiota-Gut-Brain Axis. *Physiological Reviews* (2019) doi:10.1152/physrev.00018.2018.
7. Loh, J. S. *et al.* Microbiota–gut–brain axis and its therapeutic applications in neurodegenerative diseases. *Signal Transduction and Targeted Therapy* **9**, 1–53 (2024).
8. Parker, A. *et al.* Fecal microbiota transfer between young and aged mice reverses hallmarks of the aging gut, eye, and brain. *Microbiome* **10**, 1–25 (2022).
9. D'Amato, A. *et al.* Faecal microbiota transplant from aged donor mice affects spatial learning and memory via modulating hippocampal synaptic plasticity- and neurotransmission-related proteins in young recipients. *Microbiome* **8**, 1–19 (2020).
10. Boehme, M. *et al.* Microbiota from young mice counteracts selective age-associated behavioral deficits. *Nat Aging* **1**, 666–676 (2021).
11. Wikoff, W. R. *et al.* Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proceedings of the National Academy of Sciences* **106**, 3698–3703 (2009).
12. Lai, Y. *et al.* High-coverage metabolomics uncovers microbiota-driven biochemical landscape of interorgan transport and gut-brain communication in mice. *Nat Commun* **12**, 6000 (2021).
13. Wang, Z. *et al.* Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**, 57–63 (2011).
14. Saeedi Saravi, S. S. *et al.* Gut microbiota-dependent increase in phenylacetic acid induces

1073 endothelial cell senescence during aging. *Nat Aging* **5**, 1025–1045 (2025).
1074 15. Knox, E. G., Aburto, M. R., Clarke, G., Cryan, J. F. & O'Driscoll, C. M. The blood-brain
1075 barrier in aging and neurodegeneration. *Molecular Psychiatry* **27**, 2659–2673 (2022).
1076 16. Shi, S. M. *et al.* Glycocalyx dysregulation impairs blood–brain barrier in ageing and
1077 disease. *Nature* **639**, 985–994 (2025).
1078 17. Panyard, D. J., Yu, B. & Snyder, M. P. The metabolomics of human aging: Advances,
1079 challenges, and opportunities. *Sci Adv* **8**, eadd6155 (2022).
1080 18. Menni, C. *et al.* Metabolomic markers reveal novel pathways of ageing and early
1081 development in human populations. *Int J Epidemiol* **42**, 1111–1119 (2013).
1082 19. Dunn, W. B. *et al.* Molecular phenotyping of a UK population: defining the human serum
1083 metabolome. *Metabolomics* **11**, 9–26 (2015).
1084 20. Darst, B. F., Koscik, R. L., Hogan, K. J., Johnson, S. C. & Engelman, C. D. Longitudinal
1085 plasma metabolomics of aging and sex. *Aging (Albany NY)* **11**, 1262–1282 (2019).
1086 21. Chen, H. *et al.* Highly multiplexed bioactivity screening reveals human and microbiota
1087 metabolome-GPCRome interactions. *Cell* **186**, 3095–3110.e19 (2023).
1088 22. Dekkers, K. F. *et al.* An online atlas of human plasma metabolite signatures of gut
1089 microbiome composition. *Nat Commun* **13**, 5370 (2022).
1090 23. Bishai, J. D. & Palm, N. W. Small Molecule Metabolites at the Host-Microbiota Interface. *J*
1091 *Immunol* **207**, 1725–1733 (2021).
1092 24. Bröer, S. *et al.* Alanine metabolism, transport, and cycling in the brain. *Journal of*
1093 *neurochemistry* **102**, (2007).
1094 25. Wishart, D. S. *et al.* HMDB 5.0: the Human Metabolome Database for 2022. *Nucleic Acids*
1095 *Res* **50**, D622–D631 (2021).
1096 26. Weksler, B., Romero, I. A. & Couraud, P.-O. The hCMEC/D3 cell line as a model of the
1097 human blood brain barrier. *Fluids Barriers CNS* **10**, 16 (2013).
1098 27. Sato, N. *et al.* Porphyromonas gingivalis lipopolysaccharide induces interleukin-6 and c-c
1099 motif chemokine ligand 2 expression in cultured hCMEC/D3 human brain microvascular
1100 endothelial cells. *Gerodontology* **39**, 139–147 (2022).
1101 28. Wilson 3rd, D. M. *et al*. Hallmarks of neurodegenerative diseases. *Cell* **186**, 693–714
1102 (2023).
1103 29. Xie, H. *et al.* The Role of Gut Microbiota-Derived Trimethylamine N-Oxide in the
1104 Pathogenesis and Treatment of Mild Cognitive Impairment. *International Journal of*
1105 *Molecular Sciences* **26**, 1373 (2025).
1106 30. MahmoudianDehkordi, S. *et al.* Altered bile acid profile associates with cognitive
1107 impairment in Alzheimer's disease-An emerging role for gut microbiome. *Alzheimers*
1108 *Dement* **15**, 76–92 (2019).
1109 31. Low, D. Y. *et al.* Diet-Related Metabolites Associated with Cognitive Decline Revealed by
1110 Untargeted Metabolomics in a Prospective Cohort. *Mol Nutr Food Res* **63**, e1900177
1111 (2019).
1112 32. Chen, S.-J. *et al.* Alteration of Gut Microbial Metabolites in the Systemic Circulation of
1113 Patients with Parkinson's Disease. *J Parkinsons Dis* **12**, 1219–1230 (2022).
1114 33. Kijpaisalratana, N. *et al.* White matter hyperintensity severity modifies gut metabolite
1115 association with cognitive outcomes. *J Prev Alzheimers Dis* **12**, 100086 (2025).
1116 34. Marksteiner, J., Blasko, I., Kemmler, G., Koal, T. & Humpel, C. Bile acid quantification of 20

1117 plasma metabolites identifies lithocholic acid as a putative biomarker in Alzheimer's
1118 disease. *Metabolomics* **14**, 1–10 (2017).
1119 35. Melo, H. M. *et al.* Palmitate Is Increased in the Cerebrospinal Fluid of Humans with Obesity
1120 and Induces Memory Impairment in Mice via Pro-inflammatory TNF-α. *Cell Rep* **30**, 2180–
1121 2194.e8 (2020).
1122 36. Lin, C.-H., Lin, Y.-N., Lane, H.-Y. & Chen, C.-J. The identification of a potential plasma
1123 metabolite marker for Alzheimer's disease by LC-MS untargeted metabolomics. *J*
1124 *Chromatogr B Analyt Technol Biomed Life Sci* **1222**, 123686 (2023).
1125 37. Kalecký, K., German, D. C., Montillo, A. A. & Bottiglieri, T. Targeted Metabolomic Analysis
1126 in Alzheimer's Disease Plasma and Brain Tissue in Non-Hispanic Whites. *Journal of*
1127 *Alzheimer's Disease* (2022).
1128 38. Chen, S.-J. *et al.* The Gut Metabolite Trimethylamine N-oxide Is Associated With
1129 Parkinson's Disease Severity and Progression. *Movement Disorders* **35**, 2115–2116
1130 (2020).
1131 39. Xu, Z. *et al.* Plasma trimethylamine N-oxide, carnitine, choline, and betaine distinguish
1132 cognition impairment in Parkinson's disease. *Parkinsonism Relat Disord* **135**, 107844
1133 (2025).
1134 40. Jonsson, A. L. *et al.* Impact of Gut Microbiota and Diet on the Development of
1135 Atherosclerosis in Apoe−/− Mice. *Arteriosclerosis, Thrombosis, and Vascular Biology* **38**,
1136 2318 (2018).
1137 41. Wei, W. *et al.* Psychological stress-induced microbial metabolite indole-3-acetate disrupts
1138 intestinal cell lineage commitment. *Cell metabolism* **36**, (2024).
1139 42. Sankowski, B. *et al.* Higher cerebrospinal fluid to plasma ratio of p-cresol sulfate and
1140 indoxyl sulfate in patients with Parkinson's disease. *Clin Chim Acta* **501**, 165–173 (2020).
1141 43. Pilitsis, J. G. *et al.* Quantification of Free Fatty Acids in Human Cerebrospinal Fluid.
1142 *Neurochemical Research* **26**, 1265–1270 (2001).
1143 44. Sarrias, M. J., Cabré, P., Martínez, E. & Artigas, F. Relationship between serotoninergic
1144 measures in blood and cerebrospinal fluid simultaneously obtained in humans. *J*
1145 *Neurochem* **54**, 783–786 (1990).
1146 45. Harnisch, L.-O. *et al.* Quantification of Bile Acids in Cerebrospinal Fluid: Results of an
1147 Observational Trial. *Biomedicines* **11**, 2947 (2023).
1148 46. Wang, J. *et al.* Peripheral serum metabolomic profiles inform central cognitive impairment.
1149 *Sci Rep* **10**, 14059 (2020).
1150 47. Bento, A. P. *et al.* An open source chemical structure curation pipeline using RDKit. *J*
1151 *Cheminform* **12**, 51 (2020).
1152 48. Djousse, L. *et al.* Serum Individual Nonesterified Fatty Acids and Risk of Heart Failure in
1153 Older Adults. *Cardiology* **146**, 351–358 (2021).
1154 49. Cui, Y. *et al.* Gas chromatography-mass spectrometry analysis of the free fatty acids in
1155 serum obtained from patients with Alzheimer's disease. *Biomed Mater Eng* **26 Suppl 1**,
1156 S2165–77 (2015).
1157 50. Dou, L. *et al.* The cardiovascular effect of the uremic solute indole-3 acetic acid. *J Am Soc*
1158 *Nephrol* **26**, 876–887 (2015).
1159 51. Lin, Y.-T. *et al.* Indole-3 acetic acid increased risk of impaired cognitive function in patients
1160 receiving hemodialysis. *Neurotoxicology* **73**, 85–91 (2019).

1161 52. Nho, K. *et al.* Altered bile acid profile in mild cognitive impairment and Alzheimer's disease:
1162 Relationship to neuroimaging and CSF biomarkers. *Alzheimers Dement* **15**, 232–244
1163 (2019).
1164 53. Bain, M. A., Faull, R., Fornasini, G., Milne, R. W. & Evans, A. M. Accumulation of
1165 trimethylamine and trimethylamine- N -oxide in end-stage renal disease patients
1166 undergoing haemodialysis. *Nephrol Dial Transplant* **21**, 1300–1304 (2006).
1167 54. Suzuki, T., Heaney, L. M., Bhandari, S. S., Jones, D. J. L. & Ng, L. L. Trimethylamine N-
1168 oxide and prognosis in acute heart failure. *Heart* **102**, 841–848 (2016).
1169 55. Stubbs, J. R. *et al.* Serum Trimethylamine-N-Oxide is Elevated in CKD and Correlates with
1170 Coronary Atherosclerosis Burden. *J Am Soc Nephrol* **27**, 305–313 (2016).
1171 56. U.S. Department of Health and Human Services, Food and Drug Administration, Center for
1172 Drug Evaluation and Research (CDER). Guidance for Industry: Estimating the Maximum
1173 Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers.
1174 *fda.gov* https://www.fda.gov/media/72309/download (2005).
1175 57. Nair, A. B. & Jacob, S. A simple practice guide for dose conversion between animals and
1176 human. *J Basic Clin Pharm* **7**, 27–31 (2016).
1177 58. Hoyles, L. *et al.* Regulation of blood-brain barrier integrity by microbiome-associated
1178 methylamines and cognition by trimethylamine N-oxide. *Microbiome* **9**, 235 (2021).
1179 59. Reinecke, F., Smeitink, J. A. M. & van der Westhuizen, F. H. OXPHOS gene expression
1180 and control in mitochondrial disorders. *Biochim Biophys Acta* **1792**, 1113–1121 (2009).
1181 60. Zhu, A., Ibrahim, J. G. & Love, M. I. Heavy-tailed prior distributions for sequence count
1182 data: removing the noise and preserving large differences. *Bioinformatics* **35**, 2084–2092
1183 (2019).
1184 61. Zuo, J. *et al.* Glycolysis Rate-Limiting Enzymes: Novel Potential Regulators of Rheumatoid
1185 Arthritis Pathogenesis. *Frontiers in Immunology* **12**, 779787 (2021).
1186 62. Zhou, R., Tardivel, A., Thorens, B., Choi, I. & Tschopp, J. Thioredoxin-interacting protein
1187 links oxidative stress to inflammasome activation. *Nat Immunol* **11**, 136–140 (2010).
1188 63. Spector, R. Trimethylamine N-oxide induces inflammation and endothelial dysfunction in
1189 human umbilical vein endothelial cells via activating ROS-TXNIP-NLRP3 inflammasome.
1190 *Biochemical and Biophysical Research Communications* **481**, 63–70 (2016).
1191 64. Li, D. *et al.* Trimethylamine-N-oxide promotes brain aging and cognitive impairment in mice.
1192 *Aging Cell* **17**, e12768 (2018).
1193 65. Ximerakis, M. *et al.* Single-cell transcriptomic profiling of the aging mouse brain. *Nat*
1194 *Neurosci* **22**, 1696–1708 (2019).
1195 66. Del Gatto-Konczak, F. *et al.* The RNA-Binding Protein TIA-1 Is a Novel Mammalian Splicing
1196 Regulator Acting through Intron Sequences Adjacent to a 5′ Splice Site. *Molecular and*
1197 *Cellular Biology* 6287–6299 (2023).
1198 67. Xu, S. *et al.* SRRM2 organizes splicing condensates to regulate alternative splicing. *Nucleic*
1199 *Acids Res* **50**, 8599–8614 (2022).
1200 68. Tagawa, T. *et al.* A virus-induced circular RNA maintains latent infection of Kaposi's
1201 sarcoma herpesvirus. *Proceedings of the National Academy of Sciences* **120**,
1202 e2212864120 (2023).
1203 69. Campagne, S. *et al.* Molecular basis of RNA-binding and autoregulation by the cancer-
1204 associated splicing factor RBM39. *Nature Communications* **14**, 1–18 (2023).

1205 70. Tollervey, J. R. *et al.* Analysis of alternative splicing associated with aging and
1206 neurodegeneration in the human brain. *Genome Res* **21**, 1572–1582 (2011).
1207 71. Wang, Y. *et al.* rMATS-turbo: an efficient and flexible computational tool for alternative
1208 splicing analysis of large-scale RNA-seq data. *Nat Protoc* **19**, 1083–1104 (2024).
1209 72. Grosz, B. R. *et al.* A novel synonymous KMT2B variant in a patient with dystonia causes
1210 aberrant splicing. *Molecular Genetics & Genomic Medicine* **10**, e1923 (2022).
1211 73. Haploinsufficiency of KMT2B, Encoding the Lysine-Specific Histone Methyltransferase 2B,
1212 Results in Early-Onset Generalized Dystonia. *The American Journal of Human Genetics*
1213 **99**, 1377–1387 (2016).
1214 74. Hu, Z.-Y., Xiao, L., Bode, A. M., Dong, Z. & Cao, Y. Glycolytic genes in cancer cells are
1215 more than glucose metabolic regulators. *J Mol Med (Berl)* **92**, 837–845 (2014).
1216 75. Fan, Z. *et al.* The vascular gene Apold1 is dispensable for normal development but controls
1217 angiogenesis under pathological conditions. *Angiogenesis* **26**, 385–407 (2023).
1218 76. Kim, K.-H., Won, J. H., Cheng, N. & Lau, L. F. The matricellular protein CCN1 in tissue
1219 injury repair. *J Cell Commun Signal* **12**, 273–279 (2018).
1220 77. Shi, H. *et al.* Kruppel-like factor 2 protects against ischemic stroke by regulating endothelial
1221 blood brain barrier function. *Am J Physiol Heart Circ Physiol* **304**, H796–805 (2013).
1222 78. Hatipoglu, O. F. *et al.* ADAMTS1 is a unique hypoxic early response gene expressed by
1223 endothelial cells. *J Biol Chem* **284**, 16325–16333 (2009).
1224 79. Seo, J. H. *et al.* AKAP12 Supports Blood-Brain Barrier Integrity against Ischemic Stroke. *Int*
1225 *J Mol Sci* **21**, (2020).
1226 80. Katsuoka, F. *et al.* Genetic evidence that small maf proteins are essential for the activation
1227 of antioxidant response element-dependent genes. *Mol Cell Biol* **25**, 8044–8051 (2005).
1228 81. Smith, H. W. & Marshall, C. J. Regulation of cell signalling by uPAR. *Nature Reviews*
1229 *Molecular Cell Biology* **11**, 23–36 (2010).
1230 82. Foltyn, M. *et al.* The physiological mTOR complex 1 inhibitor DDIT4 mediates therapy
1231 resistance in glioblastoma. *Br J Cancer* **120**, 481–487 (2019).
1232 83. Hai, T., Wolfgang, C. D., Marsee, D. K., Allen, A. E. & Sivaprasad, U. ATF3 and stress
1233 responses. *Gene Expr* **7**, 321–335 (1999).
1234 84. Yuniati, L., Scheijen, B., van der Meer, L. T. & van Leeuwen, F. N. Tumor suppressors
1235 BTG1 and BTG2: Beyond growth control. *Journal of Cellular Physiology* **234**, 5379–5389
1236 (2019).
1237 85. Huang, T., Yin, J., Ren, S. 'e & Zhang, X. Protective effects of KLF4 on blood-brain barrier
1238 and oxidative stress after cerebral ischemia-reperfusion in rats through the Nrf2/Trx1
1239 pathway. *Cytokine* **169**, 156288 (2023).
1240 86. Sangwung, P. *et al.* KLF2 and KLF4 control endothelial identity and vascular integrity. *JCI*
1241 *Insight* **2**, e91700 (2017).
1242 87. Sayed, N. *et al.* Clinical trial in a dish using iPSCs shows lovastatin improves endothelial
1243 dysfunction and cellular cross-talk in LMNA cardiomyopathy. *Sci Transl Med* **12**, (2020).
1244 88. Xu, D. & Li, C. Gene 33/Mig6/ERRFI1, an Adapter Protein with Complex Functions in Cell
1245 Biology and Human Diseases. *Cells* **10**, 1574 (2021).
1246 89. Fritz, G. & Kaina, B. Transcriptional activation of the small GTPase gene rhoB by genotoxic
1247 stress is regulated via a CCAAT element. *Nucleic Acids Res* **29**, 792–798 (2001).
1248 90. Luca, F. *et al.* Adaptive Variation Regulates the Expression of the Human SGK1 Gene in

Response to Stress. *PLOS Genetics* **5**, e1000489 (2009).
91. Meng, F. *et al.* Genomic editing of intronic enhancers unveils their role in fine-tuning tissue-specific gene expression in Arabidopsis thaliana. *Plant Cell* **33**, 1997–2014 (2021).
92. ENCODE Project Consortium *et al.* Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699–710 (2020).
93. Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods* **14**, 975–978 (2017).
94. Rauluseviciute, I. *et al.* JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* **52**, D174–D182 (2024).
95. Jin, W. *et al*. Microbiota-derived bile acids antagonize the host androgen receptor and drive anti-tumor immunity. *Cell* **188**, 2336–2353.e38 (2025).
96. Klemm, S. L., Shipony, Z. & Greenleaf, W. J. Chromatin accessibility and the regulatory epigenome. *Nature Reviews Genetics* **20**, 207–220 (2019).
97. Janowsky, J. S. The role of androgens in cognition and brain aging in men. *Neuroscience* **138**, 1015–1020 (2006).
98. Jardí, F. *et al.* Testosterone boosts physical activity in male mice via dopaminergic pathways. *Sci Rep* **8**, 957 (2018).
99. Davey, R. A. *et al.* Androgen Action via the Androgen Receptor in Neurons Within the Brain Positively Regulates Muscle Mass in Male Mice. *Endocrinology* **158**, 3684–3695 (2017).
100. Gräff, J. & Tsai, L.-H. The potential of HDAC inhibitors as cognitive enhancers. *Annu Rev Pharmacol Toxicol* **53**, 311–330 (2013).
101. Zhao, S. *et al.* Post sleeve gastrectomy-enriched gut commensal Clostridia promotes secondary bile acid increase and weight loss. *Gut Microbes* **17**, 2462261 (2025).
102. Qi, X. *et al.* Gut microbiota–bile acid–interleukin-22 axis orchestrates polycystic ovary syndrome. *Nature Medicine* **25**, 1225–1233 (2019).
103. Meessen, E. C. E. *et al.* Glycodeoxycholic Acid Inhibits Primary Bile Acid Synthesis With Minor Effects on Glucose and Lipid Homeostasis in Humans. *J Clin Endocrinol Metab* **110**, 1468–1477 (2025).
104. Monteiro-Cardoso, V. F., Corlianò, M. & Singaraja, R. R. Bile Acids: A Communication Channel in the Gut-Brain Axis. *Neuromolecular Med* **23**, 99–117 (2021).
105. Taylor, T. N., Greene, J. G. & Miller, G. W. Behavioral phenotyping of mouse models of Parkinson's disease. *Behavioural Brain Research* **211**, 1–10 (2010).
106. Yalçin, M., Grande, V., Outeiro, T. F. & Relógio, A. Circadian clock dysfunction in Parkinson's disease: mechanisms, consequences, and therapeutic strategy. *NPJ Parkinsons Dis* **11**, 213 (2025).
107. Kirshenbaum, G. S., Dachtler, J., Roder, J. C. & Clapcote, S. J. Characterization of cognitive deficits in mice with an alternating hemiplegia-linked mutation. *Behav Neurosci* **129**, 822–831 (2015).
108. Heredia-López, F. J. *et al.* An automated Y-maze based on a reduced instruction set computer (RISC) microcontroller for the assessment of continuous spontaneous alternation in rats. *Behav Res Methods* **48**, 1631–1643 (2016).
109. Bendl, J. *et al.* The three-dimensional landscape of cortical chromatin accessibility in Alzheimer's disease. *Nat Neurosci* **25**, 1366–1378 (2022).

110. Li, X. *et al.* Transcriptional and epigenetic decoding of the microglial aging process. *Nature Aging* **3**, 1288–1311 (2023).
111. Canyelles, M. *et al.* Gut Microbiota-Derived TMAO: A Causal Factor Promoting Atherosclerotic Cardiovascular Disease? *Int J Mol Sci* **24**, (2023).
112. Stilling, R. M. *et al.* Social interaction-induced activation of RNA splicing in the amygdala of microbiome-deficient mice. *Elife* **7**, (2018).
113. Thion, M. S. *et al*. Microbiome Influences Prenatal and Adult Microglia in a Sex-Specific Manner. *Cell* **172**, 500–516.e16 (2018).
114. Kaur, H., Singh, Y., Singh, S. & Singh, R. B. Gut microbiome-mediated epigenetic regulation of brain disorder and application of machine learning for multi-omics data analysis. *Genome* **64**, 355–371 (2021).
115. Molinaro, A. *et al.* Imidazole propionate is increased in diabetes and associated with dietary patterns and altered microbial ecology. *Nat Commun* **11**, 5881 (2020).
116. Shah, S. N. *et al.* Cerebrovascular damage caused by the gut microbe/host co-metabolite p-cresol sulfate is prevented by blockade of the EGF receptor. *Gut Microbes* **16**, 2431651 (2024).
117. Ticinesi, A., Guerra, A., Nouvenne, A., Meschi, T. & Maggi, S. Disentangling the Complexity of Nutrition, Frailty and Gut Microbial Pathways during Aging: A Focus on Hippuric Acid. *Nutrients* **15**, (2023).
118. Li, M. *et al.* Gut microbiota metabolite indole-3-acetic acid maintains intestinal epithelial homeostasis through mucin sulfation. *Gut microbes* **16**, (2024).
119. Mayhan, W. G. VEGF increases permeability of the blood-brain barrier via a nitric oxide synthase/cGMP-dependent pathway. *Am J Physiol* **276**, C1148–53 (1999).
120. Qing, Y. *et al.* Targeted metabolomics reveals aberrant profiles of serum bile acids in patients with schizophrenia. *Schizophrenia (Heidelb)* **8**, 65 (2022).
121. Pavlova, T. *et al.* Urinary intermediates of tryptophan as indicators of the gut microbial metabolism. *Anal Chim Acta* **987**, 72–80 (2017).
122. Yang, B. *et al.* HIV-Differentiated Metabolite N-Acetyl-L-Alanine Dysregulates Human Natural Killer Cell Responses to Mycobacterium tuberculosis Infection. *International Journal of Molecular Sciences* **24**, 7267 (2023).
123. Green, M. R. & Sambrook, J. Precipitation of RNA with Ethanol. *Cold Spring Harb Protoc* **2020**, 101717 (2020).
124. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* **34**, 525–527 (2016).
125. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
126. Raudvere, U. *et al.* g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res* **47**, W191–W198 (2019).
127. Garrido-Martín, D., Palumbo, E., Guigó, R. & Breschi, A. ggsashimi: Sashimi plot revised for browser- and annotation-independent splicing visualization. *PLoS Comput Biol* **14**, e1006360 (2018).
128. Liu, B. B. et al. Dissecting regulatory syntax in human development with scalable multiomics and deep learning. Preprint at https://www.biorxiv.org/content/10.1101/2025.04.30.651381v1 (2025).

129. Marinov, G. K., Shipony, Z., Kundaje, A. & Greenleaf, W. J. Genome-Wide Mapping of Active Regulatory Elements Using ATAC-seq. *Methods Mol Biol* **2611**, 3–19 (2023).
130. Picelli, S. *et al.* Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res* **24**, 2033–2040 (2014).
131. Marinov, G. K. & Shipony, Z. Interrogating the Accessible Chromatin Landscape of Eukaryote Genomes Using ATAC-seq. *Methods Mol Biol* **2243**, 183–226 (2021).
132. Grandi, F. C., Modi, H., Kampman, L. & Corces, M. R. Chromatin accessibility profiling by ATAC-seq. *Nat Protoc* **17**, 1518–1552 (2022).
133. Systematic identification of age-related microbiome metabolites that impact the brain. National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1328687 (2025).