Epigenetic MRI: Noninvasive Imaging of DNA Methylation in the Brain

Fan Lam[1,2,3,4+], James Chu[2,4+], Ji Sun Choi[2,4], Chang Cao[2,3], T. Kevin Hitchens[5], Scott K. Silverman[6], Zhi-Pei Liang[1,2,7], Ryan N. Dilger[2,3,8,9], Gene E. Robinson[3,4,9,10*] and King C. Li[2,4*]

+ Equal contributions

* Corresponding authors

1 Department of Bioengineering, University of Illinois Urbana-Champaign

2 Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign

3 Carl R. Woese Institute for Genomic Biology, University of Illinois Urbana-Champaign

4 Carle-Illinois College of Medicine, University of Illinois Urbana-Champaign

5 Department of Neurobiology and Animal Imaging Center, University of Pittsburgh

6 Department of Chemistry, University of Illinois Urbana-Champaign

7 Department of Electrical and Computer Engineering, University of Illinois Urbana-Champaign

8 Department of Animal Sciences, University of Illinois Urbana-Champaign

9 Neuroscience Program, University of Illinois Urbana-Champaign

10 Department of Entomology, University of Illinois Urbana-Champaign

Correspondence to

Gene E. Robinson (generobi@illinois.edu) and King C. Li (kingli@illinois.edu).

Abstract

It has been recognized that an integration of neuronal and genetic mechanisms supports brain function, regulates behaviour, and underpins response to environmental or disease stimuli. Several different technologies are available for imaging and studying neuronal activity in living brains, such as functional magnetic resonance imaging (fMRI), and have been translated to humans. However, the tools available to measure gene expression are destructive. Here we present a method, called epigenetic MRI (eMRI), that overcomes this limitation. eMRI achieves for the first time direct and noninvasive imaging of DNA methylation, a major gene expression regulator, in intact brains. eMRI exploits the methionine metabolic pathways that are responsible for DNA methylation to label the methyl-cytosine in brain genomic DNA through carbon-13 enriched diets. It then uses a novel carbon-13 magnetic resonance spectroscopic imaging (^{13}C-MRSI) method to map the spatial distribution of labeled DNA. We demonstrated successful ^{13}C labeling of brain DNA through diet using mass spectrometry, and robust and specific detection of labeled DNA using ^{13}C-MRSI. We used eMRI and a biomedical piglet model to produce the first DNA methylation map of an intact brain hemisphere. With both noninvasive labeling and imaging, we expect eMRI to be readily translated to humans and thus enable many new investigations into the epigenetic basis of brain function, behavior, and disease.

The brain is highly dynamic, ever-changing in its structure and function as a result of development, aging, environmental influence, and disease. Two fundamental mechanisms underpin these changes: neuronal activation, which occurs over relatively short time scales (milliseconds, seconds, and minutes), and gene expression, which occurs over longer time scales (hours, days, or even longer) [1–3]. The development of imaging technologies in the past decades have transformed how we investigate these mechanisms.

Functional magnetic resonance imaging (fMRI), in particular, has revolutionized our understanding of the human brain by providing a powerful non-destructive method to image some of the physiological correlates of changes in neural activity, e.g., changes in blood flow and oxygen metabolism [4–7]. By contrast, the technologies to image changes in gene expression have been limited to destructive methods that require invasive sampling and processing of brain tissues [8–11]. Although these techniques have provided tremendous knowledge about the dynamic aspects of gene expression and gene regulation in the brain, especially in animal models, their destructive nature makes longitudinal studies of the same samples impossible, thus fundamentally limiting our ability to translate and expand scientific discoveries to human brains. This is especially unfortunate because these longer-term changes in brain function have been shown to play critical roles in both brain diseases and responses of the brain to environmental change [1-3]. Therefore, the capability to measure, map, and visualize epigenetic modifications that regulate gene expression in the brain noninvasively, especially in humans, will revolutionize the study of brain function, behavior, and disease [12].

The quest to map brain gene expression and regulation in living organisms is not new. Significant efforts have been made to image reporter genes or associated enzymes using optical techniques, PET, or MRI [13–17]. However, these methods are either limited to model organisms or require transgenic animals engineered to express a particular reporter gene and an exogenous contrast probe that interacts with the reporter gene to produce the desired images [15,16]. Therefore, they do not offer a clear path for translation to humans. Furthermore, these methods are limited to just a few genes, and therefore cannot provide a comprehensive portrait of gene expression.

Noninvasive PET imaging of histone deacetylases (HDACs) in the human brain was recently demonstrated, using a radioactive tracer that can pass the blood-brain barrier (BBB) and bind specifically to HDAC isoforms [18,19]. This provides a means to probe a major form of epigenetic gene regulation, histone acetylation, in the human brain. However, it requires introducing radioactive materials into the body, and only targets one of the enzymes that regulate histone acetylation rather than histone acetylation itself. Moreover, PET lacks the specificity to distinguish between the target molecule and downstream metabolic products [20–22]. Other imaging epigenetics approaches have been used, including studies investigating correlations between structural and functional brain imaging features revealed by MRI and gene expression or DNA methylation patterns measured from peripheral tissues such as saliva or blood [23–28]. These studies have revealed important results, but they can only provide indirect insights into brain gene expression and epigenetic regulation.

We present here the first successful imaging of brain DNA methylation using a novel approach we call “epigenetic MRI” (eMRI), which integrates stable isotope 5-methyl-2’-deoxycytidine (5mdC) labeling and magnetic resonance spectroscopic imaging (MRSI). Using a biomedical piglet model, a customized diet enriched in ^{13}C-methionine (^{13}C-Met), and innovations to ^{13}C-MRSI, we report robust DNA methylation mapping in intact brain hemispheres. Given the noninvasive nature of our labeling and imaging approach, these results provide a clear path towards a noninvasive global DNA methylation imaging paradigm for living human brains. Because DNA methylation is one of the major regulators of gene expression in mammals [29], the eMRI method promises to transform our understanding of the brain.

In vivo brain DNA methylation labeling using ^{13}C-methionine enriched diet

A key feature of eMRI is the use of a natural amino acid involved in DNA methylation that is labeled with an NMR-active stable isotope, making methylated DNA detectable with noninvasive imaging methods (Fig. 1). We chose ^{13}C-Met because Met is an essential amino acid obtained from the diet and the major methyl group donor for DNA methylation. Met is also already commonly used as a nutritional supplement, including for humans. We designed a special diet that replaced all protein with free amino acids in the proper proportions for a biomedical piglet model. We then substituted all Met

with enriched ^{13}C-Met (Fig. 1a). Ingested ^{13}C-Met crosses the BBB [30–32] and is converted into ^{13}C-*S*-adenosylmethionine (^{13}C-SAM), which then provides the ^{13}C-methyl group for cytosine methylation, effectively labeling the 5mdC in brain genomic DNA (Fig. 1b). This is similar to the isotopic labeling scheme used to study the conversion of 5mdC to 5hmdC (5-hydroxymethyl-2'-deoxycytidine) in laboratory mice with triply deuterated ^{13}C-Met followed by mass spectrometry [33]. After a diet replacement period of either 10 or 32 days, the piglets were sacrificed. Following brain dissection, one hemisphere was used for tissue sampling from various brain regions, DNA extraction, and analysis by LC-MS/MS to confirm DNA labeling; the other intact hemisphere was used for noninvasive imaging by ^{13}C NMR and MRSI analysis (Fig. 1c). Control piglets were fed with the same formula for 10 and 32 days, but all ^{13}C-Met was replaced with regular non-isotope-labeled Met.

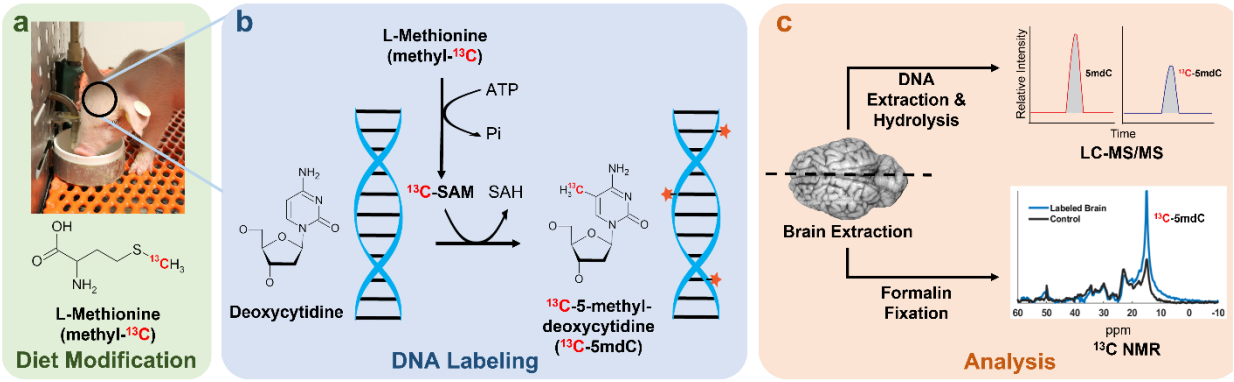

Fig. 1 Overall framework of dietary ^{13}C isotopic labeling of genomic DNA in the brain and epigenetic magnetic resonance imaging (eMRI). a. Specialized diet with enriched ^{13}C-Met was ingested by neonatal piglets (through a milk replacement formula) for either 10 or 32 days. Age-matched controls took the same diet, but with all ^{13}C-Met replaced by regular Met. **b**. The ingested ^{13}C-Met passes through the BBB and is converted to ^{13}C-SAM, which is the principal methyl group donor for DNA methylation, producing ^{13}C-methyl-labeled DNA. **c**. Brain tissues were sampled from one hemisphere for DNA extraction and LC-MS/MS analysis to confirm labeling. The other intact hemisphere was analyzed using ^{13}C NMR and MRSI to demonstrate the feasibility of noninvasive detection and imaging of ^{13}C labeled DNA.

LC-MS/MS analysis revealed effective labeling of brain 5mdC after feeding for 10 days with the ^{13}C-Met enriched diet. Relative to the control piglets, we detected a 3-4% increase in ^{13}C-labeled 5mdC (as a fraction of total 5mdC) across eight different brain

regions sampled (Fig. 2a). Feeding for 32 days with the ^{13}C-Met enriched diet resulted in much stronger labeling, with >12% increase in ^{13}C-5mdC compared to the age-matched controls (Fig. 2b). The baseline ^{13}C-labeled 5mdC signal was ~5% in the controls at both ages (Fig. 2c), which is consistent with the 1.1% natural abundance of the ^{13}C isotope and accounting for the five carbons in 5mdC [34]. The 32-day labeled samples produced >2× labeling percentage compared to the 10-day labeled samples (Fig. 2d). As expected, the baseline ^{13}C-5mdC percentage did not differ among the eight brain regions or between the control piglets fed for either 10 and 32 days (Fig. 2c, e). There are differences in ^{13}C-5mdC percentage across brain regions in the labeled samples, but these regional differences were not statistically significant for either time point (Fig. 2e, one-way ANOVA followed by Tukey's multiple comparison test). Representative LC-MS/MS data revealing the ^{13}C-5mdC and 5mdC peaks are shown in Fig. S1.

Employing the two different feeding times enabled us to make two interesting observations about DNA methylation dynamics in the brain. First, there was a nonlinear increase in labeling from day 10 to 32, much greater than from day 0 to 10 (Fig. 2d). This increase occurred despite a relative decrease in daily ^{13}C-Met intake (in terms of mg/kg body weight/day) due to rapid increases in body weight over the same time period. We speculate this to be the result of an increase in DNA methylation events and the methylation-demethylation turnover cycle as part of normal brain development during the second period. Second, the 2-fold increase in ^{13}C-5mdC labeling percentage detected in piglets fed ^{13}C-Met for 32 days relative to 10 days vastly outpaced the ~20% increase in brain volume that occurred during the same period (Fig. 2f). This finding suggests that increases in tissue volume and/or neuronal density during brain development are not the only drivers of increased ^{13}C incorporation into brain DNA. This implicates either intrinsic age-related changes in methylation dynamics or extrinsic processes such as learning and memory formation, which are known to affect DNA methylation in the brain [1–3,35–37]. They may cause unlabeled methyl groups to be replaced with ^{13}C-labeled methyl groups through turnover or cause unmethylated cytosine to be methylated, to partially explain the observed large increase in labeling across the brain.

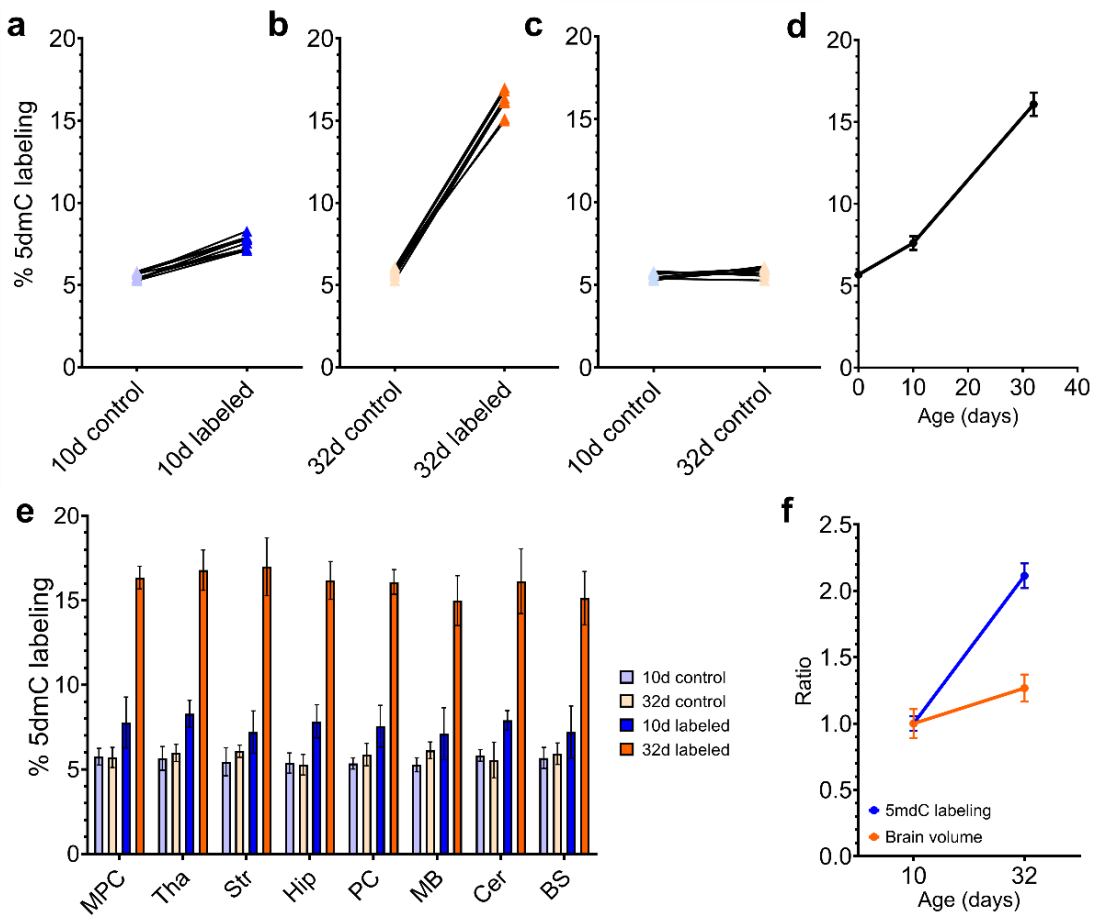

Figure 2: Dietary ^{13}C labeling of genomic DNA in piglet brains confirmed by LC-MS/MS. **a**. After 10 days of feeding on the ^{13}C-Met enriched diet, ^{13}C-5mdC in the brain DNA (ratio between ^{13}C-5mdC and total 5mdC) increased from ~5% to 8-9%, with no statistically significant differences across the eight sampled brain regions. **b**. 32 days of enriched diet significantly increased ^{13}C-5mdC level to ~15-18%. **c**. The age-matched controls for both 10-day and 32-day diets have ~5% ^{13}C-5mdC, consistent with the 1.1% natural abundance of the ^{13}C isotope (accounting for the five carbons in 5mdC) and validating our LC-MS/MS analysis. **d**. Since 5% ^{13}C-5mdC can be assumed at birth, we back-extrapolated the curve for percentage versus days of diet. A nonlinear increase in labeling can be observed, with the increase from day 10 to 32 much greater than for day 0 to 10. **e**. A more comprehensive comparison of labeling percentages for different diet groups and brain regions is shown. The error bars indicate standard deviations. No regional differences were observed for the controls. **f**. The ^{13}C-5mdC increase from 10 to 32 days with ^{13}C diet (~2×) is considerably more than the brain volume increase that occurred over the same time period (~20%). Our calibration curve with MS peak areas versus concentrations of 5mdC confirms that the range reported here is

within the linear range of our LC-MS/MS protocol (data not shown). Brain region definitions: medial prefrontal cortex (MPC), thalamus (Tha), striatum (Str), Hip (hippocampus), perirhinal cortex (PC), midbrain (MB), cerebellum (Cer), brain stem (BS).

Noninvasive detection of labeled methylated DNA using ^{13}C NMR

The strong isotope labeling of methylated DNA in the brain enables the possibility of using noninvasive stable isotopic detection methods. We chose ^{13}C NMR to detect and quantify the ^{13}C-labeled 5mdC within the genomic DNA of the piglet brain. ^{13}C NMR and imaging using isotope-enriched molecules has been used in various basic biological studies and clinical applications [38–41]. It offers several unique features as a tool of choice to image DNA methylation. First, its nondestructive and nonradioactive nature presents strong potential for *in vivo* translation to humans. Second, the broad chemical shift dispersion allows for detecting signals specifically from ^{13}C-5mdC in genomic DNA relative to ^{13}C signals from other brain molecules.

We first performed a proof-of-principle *in vitro* experiment using synthetic DNA oligonucleotides with a well-defined number of 5mdC in the sequence. Two results from the ^{13}C NMR data obtained from these oligonucleotides demonstrated the feasibility of this approach (Fig. S2). First, we were able to specifically detect the ^{13}C NMR signal from the methyl groups on 5mdC in the DNA. Second, we were able to identify the chemical shift of interest for the ^{13}C-5mdC (~15 ppm) for the subsequent brain experiments. More generally, the use of stable isotope labeling offers a unique window to probe dynamic changes in DNA methylation longitudinally over long time scales, relevant to processes related to aging, neurological disease, learning, and biological embedding of environmental stimuli. This would otherwise be difficult to achieve with less stable isotopes that are typically used to study faster metabolic processes.

A fundamental technical challenge for *in vivo* ^{13}C NMR is low sensitivity, e.g., ~60× lower than the commonly used ^{1}H NMR. While hyperpolarized-^{13}C (HP-^{13}C) has been extensively explored in recent years for *in vivo* ^{13}C NMR and imaging because of its ability to dramatically boost signal strength [42], it is not suitable for methylation labeling and imaging due to the rapid loss of spin polarization in the labeled substrate after its introduction into subjects. DNA methylation labeling is a slower process than labeling

associated with the metabolism of glucose or pyruvate which are typically utilized for HP-^{13}C MRI [43]. We overcame this challenge by integrating effective labeling using enriched ^{13}C-Met, ultrahigh-field imaging systems, and advanced signal processing techniques. We imaged the ^{13}C-labeled piglet brains using a 11.7 T microimaging system (Figure 3) and first obtained whole-sample ^{13}C NMR spectra from 8 brains (2 labeled for 10 days, 2 labeled for 32 days, and 2 control unlabeled brains at each of 10 and 32 days). Representative sample spectra are shown in Fig. 3c. Consistent results were obtained for piglets within the four groups (Fig. 3d for 10-day control and labeled; Fig. 3e for 32-day control and labeled). A significant signal difference attributed to ^{13}C-5mdC (at around 15 ppm in the spectra, as expected based on the data from the synthetic oligonucleotides), was observed between the control and labeled brains, with the 32-day labeled brain exhibiting significantly stronger signal than the 10-day labeled and control brains (Fig. 3f). The whole-sample spectra were parametrically fitted using an in-house method to quantify the amount of ^{13}C NMR 5mdC signal (Fig. S3). The signal increase from 10-day to 32-day labeling matches well with the increase in labeling percentage measured by LC-MS/MS analysis, and again significantly outpaced brain volume growth (Fig. 3g). These results confirm that increased labeling and the resulting signals are not driven solely by increases in brain volume and cell density.

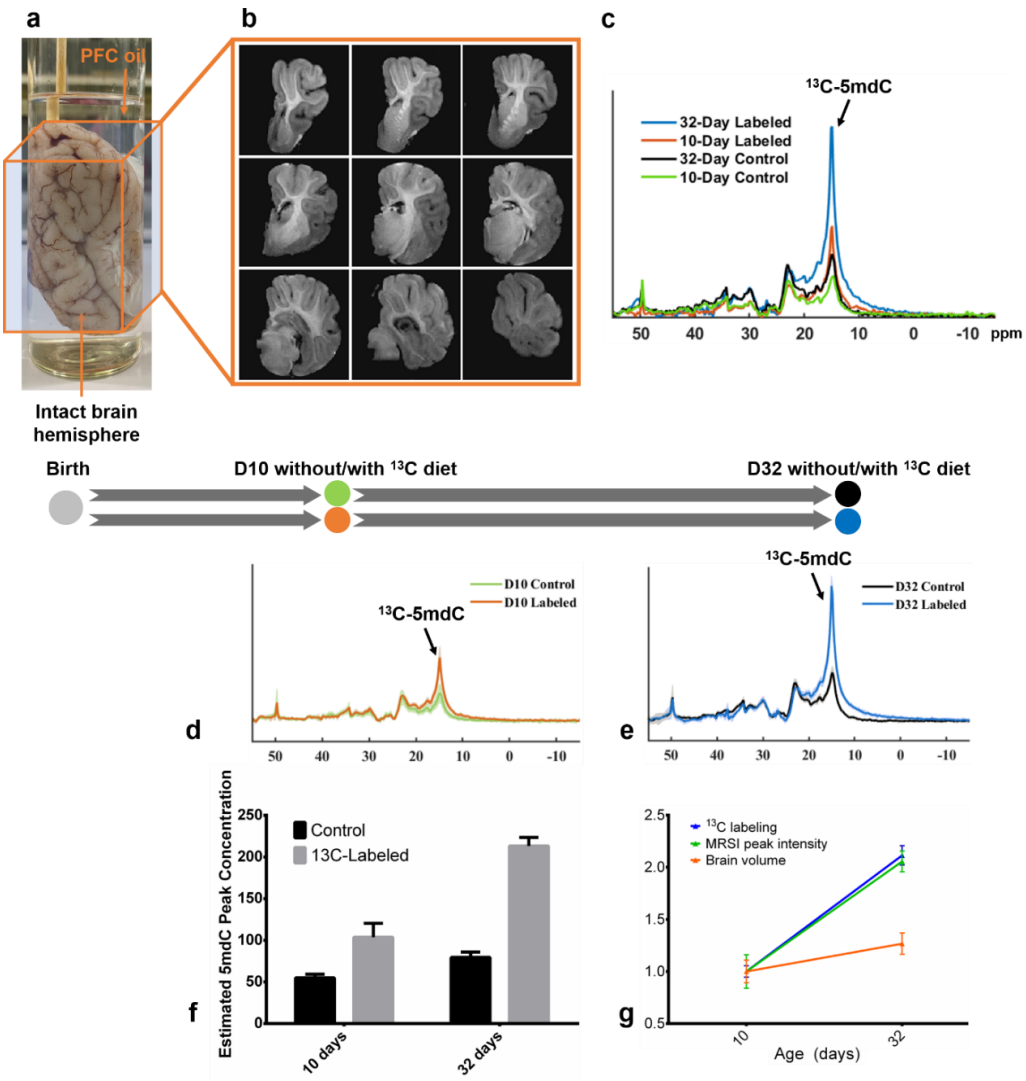

Fig. 3 ^{13}C NMR detects labeled DNA and age-dependent labeling changes in the brain. **a**. Illustration of the sample setup for spectroscopy and imaging experiments. The intact brain hemisphere was submerged in PFC oil for susceptibility matching to improve magnetic field homogeneity. **b**. A 3D T1-weighted anatomical MR image from one of the brain samples. **c**. Whole-sample ^{13}C NMR spectra acquired from different brains at different time points and dietary conditions, i.e., 10-day control diet (green), 10-day ^{13}C-Met diet (orange), 32-day control diet (black), and 32-day ^{13}C-Met diet (blue). Data for each sample were acquired in 9 h experiments. **d** and **e**. A total of eight brain hemispheres were measured, four for each time point and two for each dietary condition among the four (without and with ^{13}C-Met diet). Consistent results were obtained from each group. Labeled samples (10 days in panel d or 32 days in panel e) produced significantly stronger ^{13}C-5mdC signals (at around 15 ppm in the spectra, as marked with arrows) compared to age-matched controls. **f**. Brains labeled for 32 days exhibited significantly stronger signals than those labeled for 10 days. The NMR signal increase for the controls from 10 to 32 days was due to brain growth and is attributed to natural abundance ^{13}C NMR signals from 5mdC and methyl groups on thymidine, with the latter a

larger contributor due to higher relative abundance. Nevertheless, the 10-day labeled sample still showed a stronger signal than the 32-day control, indicating the signals measured are primarily from ^{13}C-5mdC labeling. **g**. The signal increase from 10-day to 32-day labeled brains approximately matches the labeling percentage increment measured by LC-MS/MS (Fig. 2f) and outpaced brain volume growth.

Imaging brain DNA methylation in piglets using ^{13}C-MRSI

The ability of ^{13}C NMR spectroscopy to noninvasively detect ^{13}C-5mdC signals from genomic DNA establishes the premise of using MRSI to image brain DNA methylation *in vivo* by mapping ^{13}C-labeled 5mdC. Even though DNA methylation has a strong influence on gene expression on a per-gene basis, mapping overall brain DNA methylation can provide a general measure of gene activity, analogous to how fMRI measures general neuronal activity without knowledge of specific neurons. We tested this possibility by performing ^{13}C-MRSI with a novel subspace-based processing (Methods) on our brain samples to image DNA methylation.

^{13}C-MRSI robustly mapped the spatial distribution of ^{13}C-5mdC in brain. The 32-day labeled brains, in particular, produced clearly discernible regional variations in ^{13}C-5mdC levels (Fig. 4a, b, Fig. S4). The localized ^{13}C NMR spectra extracted from several representative anatomical brain regions confirmed the regional signal variations attributed to the ^{13}C-5mdC peak around 15 ppm (Fig. 4c). Putamen (Put), caudate (Cau), and thalamus (Tha) were the regions showing the strongest signals, while the temporal and occipital lobes (TL and OL) were among the regions showing weaker signals (Fig. 4c, d). These regional measurements were confirmed to yield strong and statistically significant differences using pairwise Kruskal-Wallis with Dunn's multiple comparison tests (Fig. 4e). Reducing the acquisition time by a factor of 5 still yielded consistent regional ^{13}C-5mdC level estimates as compared to the original longer scan (Fig. S5). To validate our ^{13}C-MRSI measurements and processing methods, we performed an experiment using a physical phantom with tubes filled with ^{13}C-5mdC solutions of varying concentrations. This experiment (Fig. S6) produced concentration estimates highly consistent with the true concentrations in the phantom (correlation >0.99), confirming the robustness of our ^{13}C-MRSI method.

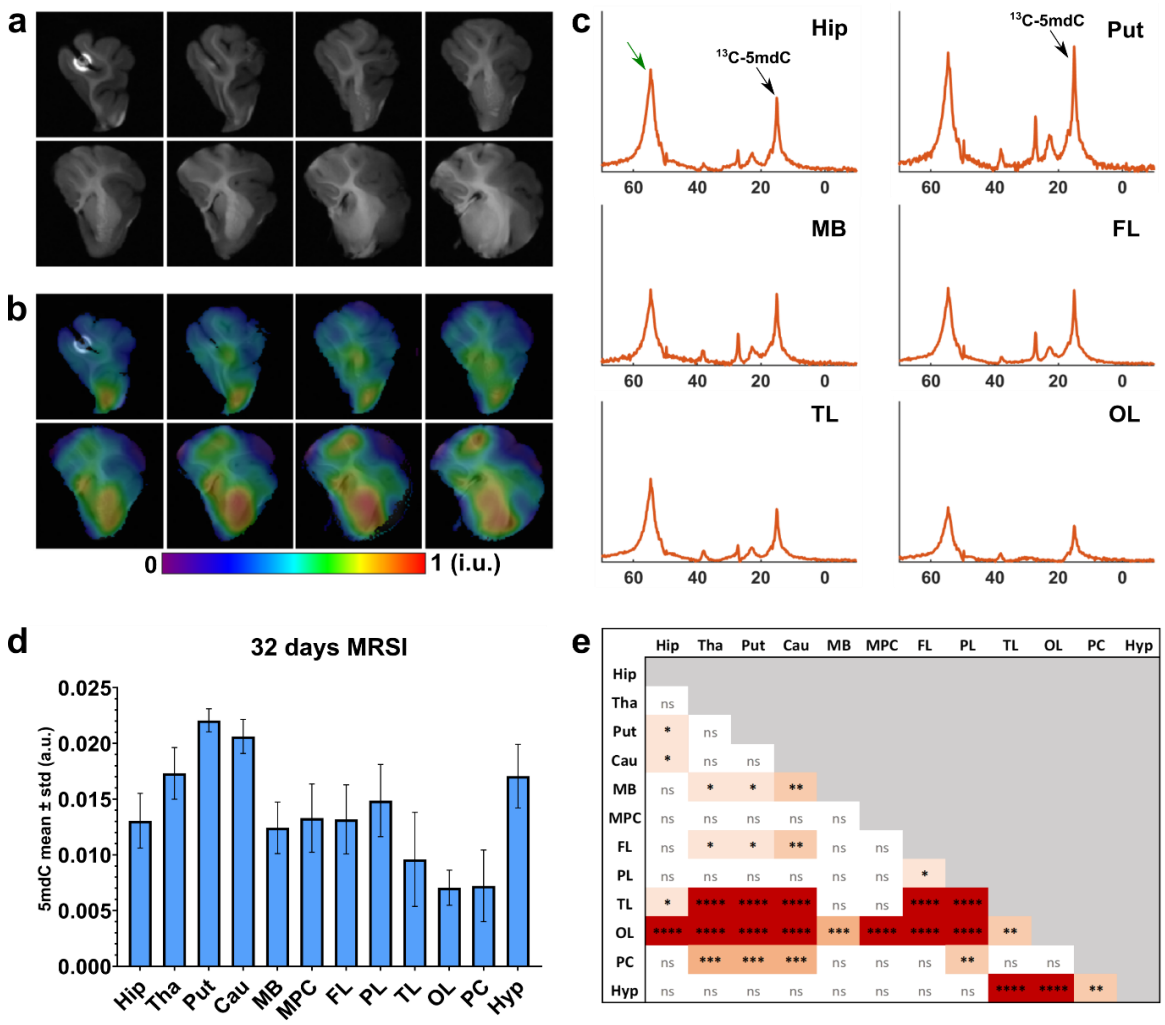

Fig. 4 Mapping regional variations of ^{13}C-labeled DNA in the intact brain. a. Anatomical MR images corresponding to different slices across the intact 3D imaging volume. **b.** The spatial maps of labeled DNA methylation overlaid on the same image slices in a. The maps were normalized with a maximum intensity of 1 (arbitrary units). Clear spatial variations of labeled DNA methylation are observed. **c.** Localized ^{13}C NMR spectra from different brain regions (averaged for the voxels within the same region), which also exhibit clear regional variation. **d**. A quantitative comparison of ^{13}C-5mdC signal differences across brain regions, with the error bars indicating the standard deviations for individual regions. Cerebellum and part of brain stem were not included in the imaging because they did not fit into the RF probes used on the imaging system. **e**. Pairwise Kruskal-Wallis with Dunn's multiple comparison tests reveal that several regions produced significantly different ^{13}C-5mdC signals and thus different labeled methylated DNA contents. * $P < 0.05$, ** $P < 0.001$, *** $P < 0.0002$, **** $P < 0.0001$. We imaged half brain (as an intact hemisphere) due to the limitation in the sample sizes that fit in our imaging instrumentation with the available ^{13}C coil. However, this is not an inherent limitation of the imaging method and imaging the whole brain will offer stronger ^{13}C-5mdC

signals. The peak around 55 ppm indicated by the green arrow in panel c is hypothesized to be related to lipid metabolism (Fig. S3). Brain region definitions: hippocampus (Hip), thalamus (Tha), putamen (Put), caudate (Cau), midbrain (MB), medial prefrontal cortex (MPC), frontal lobe (FL), parietal lobe (PL), temporal lobe (TL), occipital lobe (OL), perirhinal cortex (PC), and hypothalamus (Hyp).

Discussion

^{13}C-Met is incorporated in metabolic pathways that lead to the methylation of various molecules in the brain in addition to DNA, e.g., RNA and proteins, potentially generating confounding signals in the ^{13}C-MRSI. However, the methyl and other alkyl groups on natural protein side chains have fairly large ^{13}C NMR chemical shift differences from the methyl group on 5mdC and should not contribute substantially to what we observe at ~15 ppm. While various types of methylated RNA may be present in the brain, most methyl groups on these molecules have sufficiently different chemical shifts from 5mdC in the DNA, except for methylated cytosines in the RNA (referred to as 5mC). The level of 5mC in mRNA is estimated at approximately 0.03–0.1% of cytosines, much rarer compared to ~4-5% of methylated cytosines in DNA [44]. Higher methylation levels in rRNA and tRNA has been reported in mouse embryonic stem cells [45], but the RNA 5mC levels in mammalian brains remain largely unknown. Furthermore, the turnover rate for RNA molecules is much more rapid than DNA [46]. Therefore, while we anticipate that the contribution from 5mC in the total RNA should not significantly affect the regional variations observed by eMRI (Fig. S7), more quantitative experiments are needed to understand this contribution. Another important signal source to consider is the methyl groups on thymidine (T), which have a very similar chemical shift as 5mdC (~1-2 ppm difference; Figure S1). However, because dT is not labeled by our inclusion of ^{13}C-Met, the confounding background signals mostly come from natural abundance T. Such signals are weak compared to those from labeling, especially for the 32-day labeled samples (Fig. 3e). For future samples with lower labeling percentage, this background signal can be removed by performing a baseline scan before labeling followed by subtraction.

An intriguing observation from comparing our LC-MS/MS and MRSI data is that the regional differences measured by eMRI through ^{13}C-MRSI are much stronger than the regional differences observed in labeling percentage (Fig. 4d vs. Fig. 2e). Our results

suggest that this difference in regional variation is not due to the mild confounding signals as discussed above. Rather, we hypothesize that the strong regional variations in the eMRI data, in contrast to the smaller variation in ^{13}C-5mdC/total 5mdC ratios, can be attributed to differences in total 5mdC content from region to region. This is supported by global methylation levels measured in tissue samples from selected regions with the most significant eMRI signal differences (i.e., putamen and temporal lobe), using an enzyme-linked immunosorbent assay (ELISA). A higher global methylation level in putamen than temporal lobe was measured, consistent with what was obtained by eMRI (Fig. S7), supporting our hypothesis. An additional possible contributing factor to the regional eMRI signal variation is regional differences in DNA density, i.e., quantity of DNA per unit mass tissue. While evidence of nonuniform cell densities in different mammalian brain regions has been reported [47,48], it is unclear how this nonuniformity translates to DNA density distributions. Experiments should be pursued to evaluate these factors more comprehensively in biomedical pig models and other species.

In analogy to how fMRI has been used as a successful surrogate for imaging global neuronal activity at whole brain and brain region levels, we propose that eMRI measurements of global DNA methylation can be used as a noninvasive surrogate for imaging global gene expression at the same levels. While this surrogate is agnostic to individual genes, we anticipate that this epigenetic neuroimaging capability will provide new molecular markers to study development, aging, disease, learning, and responses to environment.

Analysis of the publicly available Allen Human Brain Atlas (ABA) has shown regional differences in global gene expression levels (Fig. S8), which also vary with age, underlining the importance of such measures. Figure S8 also shows that regional gene expression levels from the ABA do not match the regional DNA methylation levels from the current eMRI data in any simple way, including the well-established negative relationship between gene expression and DNA promoter methylation. There are several possible reasons for this. First, the ABA data were from human adults while our eMRI data are from young piglets. Second, the canonical inverse relationship between DNA methylation and gene expression exists mainly for methylation in gene promoters, whereas eMRI measures total DNA methylation. Third, perhaps there are region-specific differences in abundance of labeled methylated RNA. These results

point to the need to conduct precise quantitative experiments on gene expression and DNA methylation to better understand the eMRI signal.

One important consideration in evaluating the potential applications of eMRI is its ability to provide information on the turnover of DNA methylation in the brain *in vivo*. We chose piglets with growing brains in our study because the ^{13}C label can be incorporated into the DNA of brain cells through both new cell formation and turnover of DNA methylation. This would increase the resulting ^{13}C labels available for MRSI detection. In adult animals and humans, brain growth is not a major factor, so incorporation of ^{13}C labels into the DNA of adult brain cells is primarily through turnover of DNA methylation [49]. No known *in vivo* technique can measure this turnover which may shed light on the possible roles of DNA methylation in encoding and retrieval of memory and other cognitive functions. Furthermore, eMRI may enable *in vivo* epigenetic study at a much longer time scale, from days to even months, which is inaccessible by radioactive labeling and invasive techniques. eMRI can be coupled with fMRI to investigate the interaction between short-term neural and long-term molecular control of brain function, to gain further insights into the regulation of behavior and brain responses to environmental or disease stimuli.

eMRI can be extended in a few different ways. Besides methylated cytosine in DNA, other molecules (such as acetylated histone proteins) that are involved in epigenetic regulation of gene expression may be isotopically labeled and detected using NMR and MRSI. Stable isotopes other than ^{13}C can be explored, e.g., using deuterated Met instead of ^{13}C-Met. The sensitivity of *in vivo* deuterium (^{2}H) NMR spectroscopy and spectroscopic imaging has been demonstrated recently using deuterated glucose for imaging oxidative metabolism and glycolysis [50,51]. However, the labeling efficiency, specificity, and sensitivity for detecting ^{2}H-labeled DNA must be carefully investigated. Additionally, ^{13}C-Met already has been approved for human use, while the biological safety of ^{2}H-Met needs further evaluation.

An important consideration for applying *in vivo* eMRI to human brains is sensitivity. While dietary administration of ^{13}C-Met leads to a high labeling percentage in neonatal piglet brains, adult human brains are expected to have a lower labeling percentage [33,49]. However, human brains have approximately 10× larger volume than the piglet brains imaged here, thus providing significantly larger quantities of 5mdC for signal

detection. Increased availability of ultrahigh-field human MRI systems (7 T or higher) provides a significant sensitivity boost. Furthermore, many potential optimizations in data acquisition (e.g., cross-polarization and weighted signal averaging) and processing (e.g., advanced denoising and reconstruction techniques) can be performed for enhancing the signal. Labeling and imaging the exchangeable pool of 5mdC may provide a unique window into human brain function.

Methods

DNA synthesis. DNA oligonucleotides were prepared by solid-phase synthesis on an ABI 394 instrument using reagents from Glen Research, including the 5-methyl-2'-deoxycytidine (5mdC) phosphoramidite. Oligonucleotides were purified by 7 M urea denaturing 20% PAGE with running buffer 1× TBE (89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3), extracted from the polyacrylamide with TEN buffer (10 mM Tris, pH 8.0, 1 mM EDTA, 300 mM NaCl), and precipitated with ethanol. The 20-nucleotide DNA sequence, designed to avoid self-dimer formation and to include C nucleotides in various sequence contexts, was 5'-CTACGCCTCGCTCGCCCCTT-3', where the ten C nucleotides were either all 5mdC or all standard dC (MW 6,081 with 5mdC). Approximately 750 nmol of each oligonucleotide was dissolved in 0.5 mL 10% D_2O/90% H_2O (1.5 mM oligonucleotide; 9 mg/mL) in a standard 5 mm NMR tube for ^{1}H and ^{13}C NMR spectroscopy.

^{13}C-MRSI phantom. ^{13}C-labeled 5mdC (5-Methyl Cytosine-^{13}C, $^{15}N_2$ Hydrochloride 99%, Cat. M294702, from Toronto Research Chemicals) was dissolved in PBS (Gibco) then serially diluted to the desired concentrations of 1 mM, 4 mM, 6 mM and 10 mM. The solutions were transferred to 10 mm SP Scienceware Thin Walled Precision NMR tubes (Wilmad-LabGlass) which were taped together to fit into the 30-mm probe of the 11.7 T microimaging system.

Piglet experiments and dietary labeling. All animal care and experimental procedures were in accordance with the National Research Council Guide for the Care and Use of Laboratory Animals and were approved by the University of Illinois Institutional Animal Care and Use Committee (Protocol #20223). Naturally farrowed, intact domestic male pigs derived from the cross of Line 2 boars and Line 3 sows (Pig Improvement Company) were obtained from a commercial swine herd and transferred to the Piglet Nutrition and Cognition Laboratory (PNCL) on post-natal day (PND) 2 (n=8). Pigs were maintained at PNCL through PND 32 per standard protocols as described elsewhere with the following modifications. All pigs were provided *ad libitum* access to water and a custom, nutritionally complete [52] milk replacer formula (TestDiet) via a semi-automated liquid delivery system that dispensed milk from 1000 to 0600 h the next day [53]. The milk replacer powder was formulated to contain 30.4% lactose, 30.3% stabilized lipids (i.e., dried fat; combination of lipid sources including sunflower oil, palm oil, and medium-chain triglycerides), and 25.4% purified amino acids; no intact protein was included in this formulation. The nutritional profile of the milk replacer powder included 0.50% methionine and 0.49% cysteine, with the control and labeling diets containing regular L-methionine (Met) and *methyl*-^{13}C-L-methionine (^{13}C-Met, 99%, Sigma-Aldrich), respectively (i.e., all Met present in the labeling diet contained the stable ^{13}C isotope label). Milk replacer formula was reconstituted fresh daily with 200 g of milk replacer powder per 800 g water. Pigs were individually housed in caging units (87.6 cm × 88.9 cm × 50.8 cm; L × W × H), which allowed each pig to see, hear, and smell, but not directly touch, adjacent pigs. A toy was provided to each pig for enrichment, and pigs were allowed approximately 15 min of direct social interaction with other pigs once daily. The ambient environment in the rearing space was maintained on a 12 h light and dark cycle from 0800 to 2000 h, with room

temperature set at 27 °C for the first 21 days of the study and gradually lowered to 22 °C during the last 7 days.

Brain biopsy and fixation. At the appropriate age (PND 10 and 32 days), each pig was humanely euthanized per standard protocol [54] and the whole brain was quickly excised post-mortem or when meeting euthanasia criteria [55]. The brain was separated into two hemispheres, with tissue samples collected from 8 different regions within the left hemisphere (medial prefrontal cortex, cerebellum, hippocampus, midbrain, brain stem, thalamus, perirhinal cortex, and striatum). Enough tissue to permit the ultimate extraction of at least 1 μg total DNA (approximately 100-250 mg) was snap-frozen on dry ice and stored at –80 °C for genomic DNA extraction. The right hemisphere from each brain remained intact (i.e., was not biopsied), and was fixed in 10% neutral-buffered formalin (NBF) at 4 °C for 72 h and stored in PBS (pH 7.4; Gibco) at 4 °C for subsequent NMR spectroscopy and imaging experiments.

Genomic DNA extraction. Pulverized tissue samples were incubated in a digestion mix (100 mM Tris, 5 mM EDTA, 200 mM NaCl, 0.2% SDS, pH 5.5) containing 400 µg ml^{-1} Proteinase K (Invitrogen) and 200 ug ml^{-1} RNase A (Invitrogen) at 50 ºC overnight. Genomic DNA was extracted with phenol:chloroform:isoamyl alcohol (25:24:1, Fisher Scientific) and precipitated using 70% ethanol (Sigma-Aldrich). Genomic DNA was resuspended in HPLC-grade water (Sigma-Aldrich) and the concentration measured with the NanoDrop Spectrophotometer (Thermo Fisher Scientific).

LC-MS/MS DNA analysis. 1 µg of genomic DNA was hydrolyzed overnight at 37 ºC with 5 U of DNA Degradase Plus (Zymo Research) following the manufacturer's protocol in 50 µl volume. Digested DNA was filtered through Amicon Ultra 10K centrifuge filters (EMD Millipore) to remove undigested polynucleotides and collected for LC-MS/MS. LC-MS/MS analysis of ^{13}C-5mdC isotope labeling was performed with a Waters SYNAPT G2Si mass spectrometer and a Waters ACQUITY UPLC H-class fitted with a Waters Cortecs C18+ column (2.1 × 150 mm, 1.6 µm particle size) at the flow rate of 150 µl min^{-1} under a gradient of ammonium acetate and acetonitrile (Fisher Scientific). Multiple reaction monitoring (MRM) was set up to capture the transitions from deoxycytidine – cytosine (228.1-112.05 Da), 5-methyldeoxycytidine to 5-methylcytosine (242.1-126.07 Da), and ^{13}C-5-methyldeoxycytidine to ^{13}C-5-methylcytosine (243.1-127.07 Da). The percentage of ^{13}C isotope labeling was determined by the following formula using peak areas of dC and 5mdC:

$$\% \ ^{13}\text{C labeling} = 100 * \frac{[^{13}C - 5mdC]}{[5mdC] + [^{13}C - 5mdC]}$$

Global DNA methylation by ELISA. Brain tissues were collected from pigs at 32 days of age (n=4; 1 female, 3 males). Putamen and temporal lobe samples (~25 mg each) were dissected from the left hemisphere of each pig and briefly homogenized with a scalpel. Samples were then flash-frozen in liquid nitrogen and stored in cryovials at –80 °C. Genomic DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen). A Nanodrop 2000 (Thermo Fisher Scientific) was used to determine DNA concentrations/yield. As a quality control check, all A_{260}/A_{280} ratio values indicated

sufficient DNA purity (values ranged from 1.87 to 2.02). Global DNA methylation was assessed via an ELISA-based commercial kit [MethylFlash Methylated DNA Quantification Kit (Fluorometric); Epigentek]. All standards and samples were run in duplicate. Absorbance was read at 450 nm using a plate reader (BioTek) and Gen5 version 3.11 software. DNA methylation levels in each brain region were calculated relative to the Methylated Control DNA standard using the single point method equation: [(Average A_{450} Sample – Average A_{450} Blank) / ((Average A_{450} Methylated Control DNA – Average A_{450} Blank) x 2)] x 100%. For calculations, the quantity of each unknown sample DNA matched the quantity of Methylated Control DNA (i.e., the blanked A_{450} value for a 100 ng DNA sample was compared to the blanked A_{450} value for 100 ng Methylated Control DNA), and the factor of 2 was to normalize the percentage of 5mC in the control DNA to 100% as suggested in the manufacturer's instructions. This calculation yields relative global methylation level with respect to the control DNA, which is sufficient for our intended regional comparison.

Statistical analysis. One-way ANOVA followed by Tukey's multiple comparison test was used to compare the ^{13}C labeling efficiencies between different brain regions as measured by LC-MS/MS. Kruskal-Wallis followed by Dunn's multiple comparison tests were used to compare the ^{13}C-MRSI signal intensities from the different brain regions of the 32-day labeled brain.

Brain NMR spectroscopy and imaging experiments. Each intact piglet brain hemisphere was placed in a 30 mm diameter glass tube containing perfluorocarbon oil (Fluorinert FC-40, Sigma-Aldrich) as a magnetic susceptibility matching fluid. A wooden applicator stick was taped in place to hold the brain tissue and prevent it from floating. All brain NMR spectroscopy and imaging experiments were performed using a Bruker AV3HD 11.7 Tesla/89 mm vertical-bore microimaging system equipped with a 16-channel shim insert and Micro2.5 gradient set with a maximum gradient of 1500 mT/m. Data were collected using a 30 mm dual-tuned $^{1}H/^{13}C$ RF resonator and ParaVision 6.0.1 (Bruker Biospin). A 200 mM ^{13}C-Met solution dissolved in PBS was used for ^{13}C calibration before the brain imaging scans.

MRI: During each experiment session, a localizer scan was first acquired, followed by field map shimming. The B0 inhomogeneity was minimized to reach a ^{1}H water linewidth of between 45 and 60 Hz. Anatomical MRI scans were acquired using axial RARE (T2 weighted) and MDEFT (T1 weighted) with the following parameters, RARE: TR/TE = 6000/5.5 ms, RARE factor = 4, number of averages (NA) = 2, field of view (FOV) = 30 × 30 mm, matrix size = 192 × 192, and 50 1 mm slices; 3D MDEFT: TR/TE = 4000/2.2 ms, inversion delay = 1050 ms, FOV = 30(x) × 30(y) × 50(z) mm^3, matrix size = 192 × 192 × 50, and 8 segments. A field map was collected covering the same FOV. Brain volume was estimated from the T1-weighted MR image for each sample. The MR image was also segmented into 12 regions of interest (ROIs), i.e., hippocampus, thalamus, putamen, caudate, midbrain, medial prefrontal cortex, frontal lobe, parietal lobe, temporal lobe, occipital lobe, perirhinal cortex and hypothalamus, for regional analysis of MRSI results. The regions for tissue sampling (for LC-MS/MS

and Fig. 2) were defined slightly differently from those for MRI-based segmentation, due to the difficulty in precisely extracting some regions from the piglet brains (e.g., putamen and caudate were assigned as striatum in the LC-MS/MS analysis). The cerebellum and brain stem were excluded in the imaging experiments because of the size limitation of the instrumentation used.

^{13}C NMR spectroscopy and MRSI: A whole-sample ^{13}C NMR spectrum was collected at 125.755 MHz using a 100 µs RF pulse centered at 20 ppm, with a 2 sec TR. Broadband ^{1}H decoupling was achieved with a WALTZ-16 decoupling scheme centered at 2 ppm (^{1}H carrier frequency). The free induction decay (FID) was collected with a 22727 Hz (~180 ppm) spectral bandwidth (BW), 1024 points, and 16238 averages for a total acquisition time of 9 h and 6 min. ^{13}C-MRSI data was acquired using a phase-encoded chemical shift imaging (CSI) sequence with a 50 µs, 30°, block RF excitation pulse, a 30 × 30 × 50 mm^3 FOV, an 8 × 8 × 8 matrix, and broadband ^{1}H decoupling described above. FIDs were collected with the same BW and number of points, TR/TE = 2000/1.16 ms and 64 averages for a total acquisition time of 18 h and 12 min. A matching ^{1}H-CSI data was collected with TR/TE of 500/1.16 ms, 16 × 16 × 8 matrix size, and 512 FID points over a 20 ppm BW (17 min), to correct for spatial misalignment between the MRI and MRSI acquisitions.

NMR and MRSI data processing. A model-based data processing method was developed to perform spectral quantification of the measured ^{13}C-MRSI data to obtain the ^{13}C-5mdC concentration maps for eMRI. The method incorporates both physics-based prior information (resonance structure of the ^{13}C NMR spectrum) and the high SNR single-voxel spectroscopy (SVS) data from the whole sample. We first performed spectral quantification of the high SNR SVS data, $s_{SV}(t)$, to derive spectral basis functions for our model. We expressed $s_{SV}(t)$ as:

$$s_{SV}(t) = h(t) \cdot \sum_{m=1}^{M} a_m e^{-t/T_{2,m} + i2\pi f_m t},$$

where a_m, f_m, and $T_{2,m}$ represent concentration, resonance (chemical shift) frequency, and spin-spin relaxation constant of the mth spectral component, respectively, and $h(t)$ accounts for non-Lorentzian spectral lineshape variations introduced by non-ideal experimental conditions (e.g., magnetic field drift, field inhomogeneity, etc.). In this work, $h(t)$ corresponds to a compensated Gaussian lineshape function so that the spectral model can be fit to the experimental data, with the fitting residual at the noise level and passing the Komolgrov Gaussianity test (Fig. S3).

After spectral quantification of the SVS data, we obtained the following spectral basis functions that contain resonance frequencies and spin-spin relaxation constants of all spectral components and the spectral lineshape compensating function:

$$\varphi_m(t) = e^{-t/T_{2,m} + i2\pi f_m t} \cdot h(t), \qquad m = 1, 2, \ldots, M.$$

To perform spectral quantification of the entire ^{13}C-MRSI data set, we first mapped the data to the spatial domain using a constrained image reconstruction method that includes B_0 field inhomogeneity correction [56]. This reconstruction step included some spatial filtering effects to enhance SNR of the reconstruction. The spectral basis functions derived from the whole-sample SVS data above were used to quantify the reconstructed spatiotemporal function $\rho(x,t)$ point by point. Considering the difference in spatial resolution, shimming condition, and sequence set up between the SVS and MRSI scans, we adjusted the spectral basis functions $\varphi_m(t)$ learned from the SVS data to better match the MRSI data. To this end, we first denoised the reconstructed spatiotemporal function $\rho(x,t)$ using a low-rank filtering method [57]. We then estimated a new spectral lineshape function $h(t)$ from the denoised MRSI data. The estimated lineshape function was applied to $\varphi_m(t)$ to obtain an improved set of spectral basis, $\{\hat{\varphi}_m(t)\}_{m=1}^{M}$. With $\{\hat{\varphi}_m(t)\}$ determined, concentrations, $\hat{c}_m(x)$, of all spectral components were determined from the original reconstructed spatiotemporal function $\rho(x,t)$ data by solving the following model fitting problem:

$$\hat{c}_m(x) = \operatorname{argmin} \left\| \rho(x,t) - \sum_{m=1}^{M} c_m(x)\hat{\varphi}_m(t) \right\|_2^2 + R(c_m(x)) \,,$$

where $R(\cdot)$ represents an edge-weighting regularization function to incorporate spatial prior information. The spatial distribution of the spectral component corresponding to ^{13}C-5mdC was separated. Regional concentrations and spectra were obtained from the 12 ROIs segmented from MRI and analyzed.

Data availability

All data generated will be made available upon reasonable request to G.E.R or K.C.L.

Code availability

Custom codes were developed to analyze the ^{13}C-NMR and ^{13}C-MRSI data. The codes will be made available upon reasonable request to G.E.R or K.C.L.

References

1. Robinson, G. E., Fernald, R. D. & Clayton, D. F. Genes and social behavior. *Science* **322**, 896 LP – 900 (2008).

2. Halder, R. *et al.* DNA methylation changes in plasticity genes accompany the formation and maintenance of memory. *Nat. Neurosci.* **19**, 102–110 (2015).

3. Traniello, I. M. & Robinson, G. E. Neural and molecular mechanisms of biological embedding of social interactions. *Annu. Rev. Neurosci.* **44**, 109–128 (2021).

4. Ogawa, S., Lee, T. M., Kay, A. R. & Tank, D. W. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc. Natl. Acad. Sci.* **87**, 9868 LP – 9872 (1990).

5. Belliveau, J. W. *et al.* Functional mapping of the human visual cortex by magnetic resonance imaging. *Science (80-.).* **254**, 716 LP – 719 (1991).

6. Kwong, K. K. *et al.* Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc. Natl. Acad. Sci.* **89**, 5675 LP – 5679 (1992).

7. Bandettini, P. A., Wong, E. C., Hinks, R. S., Tikofsky, R. S. & Hyde, J. S. Time course EPI of human brain function during task activation. *Magn. Reson. Med.* **25**, 390–397 (1992).

8. Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).

9. Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011).

10. Thompson, C. L. *et al.* A High-resolution spatiotemporal atlas of gene expression of the developing mouse brain. *Neuron* **83**, 309–323 (2014).

11. Gryzinska, M., Blaszczak, E., Strachecka, A. & Jezewska-Witkowska, G. Analysis of age-related global DNA methylation in chicken. *Biochem. Genet.* **51**, 554–563 (2013).

12. Robinson, G. E. Brains work via their genes just as much as their neurons. *The Conversation* (2015).

13. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science* **263**, (1994).

14. Contag, C. H. *et al.* Visualizing gene expression in living mammals using a bioluminescent reporter. *Photochem. Photobiol.* **66**, 523–531 (1997).

15. Louie, A. Y. *et al.* In vivo visualization of gene expression using magnetic resonance imaging. *Nat. Biotechnol.* **18**, (2000).

16. Lee, S. W., Lee, S. H. & Biswal, S. Magnetic resonance reporter gene imaging. *Theranostics* **2**, (2012).

17. Massoud, T. F. & Gambhir, S. S. Molecular imaging in living subjects: seeing fundamental biological processes in a new light. *Genes Dev.* **17**, 545–580

(2003).

18. Wey, H. Y. *et al.* Insights into neuroepigenetics through human histone deacetylase PET imaging. *Sci. Transl. Med.* **8**, (2016).

19. Gilbert, T. M. *et al.* Neuroepigenetic signatures of age and sex in the living human brain. *Nat. Commun.* **10**, (2019).

20. Vaquero, J. J. & Kinahan, P. Positron emission tomography: current challenges and opportunities for technological advances in clinical and preclinical imaging systems. *Annu. Rev. Biomed. Eng.* **17**, 385–414 (2015).

21. Hou, L. *et al.* Positron emission tomography imaging of the endocannabinoid system: opportunities and challenges in radiotracer development. *J. Med. Chem.* **64**, 123–149 (2021).

22. Lin, G. & Chung, Y. L. Current opportunities and challenges of magnetic resonance spectroscopy, positron emission tomography, and mass spectrometry imaging for mapping cancer metabolism in vivo. *Biomed Res. Int.* **2014**, (2014).

23. Dannlowski, U. *et al.* Serotonin transporter gene methylation is associated with hippocampal gray matter volume. *Hum. Brain Mapp.* **35**, 5356–5367 (2014).

24. Ursini, G. *et al.* Stress-related methylation of the Catechol-O-Methyltransferase Val158 allele predicts human prefrontal cognition and activity. *J. Neurosci.* **31**, 6692 LP – 6698 (2011).

25. Nikolova, Y. S. *et al.* Beyond genotype: serotonin transporter epigenetic modification predicts human brain function. *Nat. Neurosci.* **17**, (2014).

26. Nikolova, Y. S. & Hariri, A. R. Can we observe epigenetic effects on human brain function? *Trends Cogn. Sci.* **19**, 366–373 (2015).

27. Richiardi, J. *et al.* Correlated gene expression supports synchronous activity in brain networks. *Science.* **348**, 1241–1244 (2015).

28. Puglia, M. H., Lillard, T. S., Morris, J. P. & Connelly, J. J. Epigenetic modification of the oxytocin receptor gene influences the perception of anger and fear in the human brain. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 3308–3313 (2015).

29. Moore, L. D., Le, T. & Fan, G. DNA methylation and its basic function. *Neuropsychopharmacology* **38**, 23–38 (2013).

30. O'Tuama, L. A. *et al.* L-methionine uptake by human cerebral cortex: Maturation from infancy to old age. *J. Nucl. Med.* **32**, 16–22 (1991).

31. Valdovinos-Flores, C. & Gonsebatt, M. E. The role of amino acid transporters in GSH synthesis in the blood-brain barrier and central nervous system. *Neurochem. Int.* **61**, 405–414 (2012).

32. Zaragozá, R. Transport of amino acids across the blood-brain barrier. *Front. Physiol.* **11**, 1–11 (2020).

33. Bachman, M. *et al.* 5-Hydroxymethylcytosine is a predominantly stable DNA modification SI. *Nat. Chem.* **6**, 1049–1055 (2014).

34. Wilhelm, R., Szeitz, A., Klassen, T. L. & Mohn, W. W. Sensitive, efficient quantitation of 13C-enriched nucleic acids via ultrahigh-performance liquid chromatography-tandem mass spectrometry for applications in stable isotope probing. *Appl. Environ. Microbiol.* **80**, 7206–7211 (2014).

35. Tognini, P. *et al.* Experience-dependent DNA methylation regulates plasticity in the developing visual cortex. *Nat. Neurosci.* **18**, 956–958 (2015).

36. Oliveira, A. M. M. DNA methylation: A permissive mark in memory formation and maintenance. *Learn. Mem.* **23**, 587–593 (2016).

37. Wheater, E. N. W. *et al.* DNA methylation and brain structure and function across the life course: A systematic review. *Neurosci. Biobehav. Rev.* **113**, 133–156 (2020).

38. Li, S., Yang, J. & Shen, J. Novel strategy for cerebral 13C MRS using very low RF power for proton decoupling. *Magn. Reson. Med.* **57**, 265–271 (2007).

39. Nabuurs, C. I. H. C., Klomp, D. W. J., Veltien, A., Kan, H. E. & Heerschap, A. Localized sensitivity enhanced in vivo 13C MRS to detect glucose metabolism in the mouse brain. *Magn. Reson. Med.* **59**, 626–630 (2008).

40. Hyacinthe, J. N. *et al.* Evaluating the potential of hyperpolarised [1-13C] L-lactate as a neuroprotectant metabolic biosensor for stroke. *Sci. Rep.* **10**, 1–11 (2020).

41. Ruiz-Rodado, V., Brender, J. R., Cherukuri, M. K., Gilbert, M. R. & Larion, M. Magnetic resonance spectroscopy for the study of cns malignancies. *Prog. Nucl. Magn. Reson. Spectrosc.* **122**, 23–41 (2021).

42. Ardenkjaer-Larsen, J. H. *et al*. Increase in signal-to-noise ratio of > 10,000 times in liquid-state NMR. *Proc. Natl. Acad. Sci*. **100**, 10158 LP – 10163 (2003).

43. Ginno, P. A. *et al.* A genome-scale map of DNA methylation turnover identifies site-specific dependencies of DNMT and TET activity. *Nat. Commun.* **11**, 2680 (2020).

44. Wiener, D. & Schwartz, S. The epitranscriptome beyond m6A. *Nat. Rev. Genet.* **22**, 119–131 (2021).

45. Legrand, C. *et al*. Statistically robust methylation calling for whole-transcriptome bisulfite sequencing reveals distinct methylation patterns for mouse RNAs. *Genome Res*. **27**, 1589–1596 (2017).

46. Defoiche, J. *et al*. Measurement of ribosomal RNA turnover in vivo by use of deuterium-labeled glucose. *Clin*. *Chem*. **55**, 1824–1833 (2009).

47. Herculano-Houzel, S., Collins, C. E., Wong, P., Kaas, J. H. & Lent, R. The basic nonuniformity of the cerebral cortex. *Proc. Natl. Acad. Sci.* **105**, 12593 LP – 12598 (2008).

48. Collins, C. E., Airey, D. C., Young, N. A., Leitch, D. B. & Kaas, J. H. Neuron densities vary across and within cortical areas in primates. *Proc. Natl. Acad. Sci.* **107**, 15927 LP – 15932 (2010).

49. Lister, R. & Mukamel, E. Turning over DNA methylation in the mind. *Front*.

Neurosci. **9**, 252 (2015).

50. Lu, M., Zhu, X.-H., Zhang, Y., Mateescu, G. & Chen, W. Quantitative assessment of brain glucose metabolic rates using in vivo deuterium magnetic resonance spectroscopy. *J. Cereb. Blood Flow Metab.* **37**, 3518–3530 (2017).

51. De Feyter, H. M. *et al.* Deuterium metabolic imaging (DMI) for MRI-based 3D mapping of metabolism in vivo. *Sci. Adv.* **4**, eaat7314 (2018).

52. NRC. *Nutrient Requirements of Horses,Fifth Revised Edition, 1989*. *The National Academies Press* (1989).

53. Fil, J. E., Joung, S., Hayes, C. A. & Dilger, R. N. Influence of rearing environment on longitudinal brain development, object recognition memory, and exploratory behaviors in the domestic pig (Sus scrofa). *Front. Neurosci.* **15**, 1–16 (2021).

54. Leary, S. *et al.* AVMA guidelines for the euthanasia of animals: 2020 Edition*. in *AVMA Guidelines for the Euthanasia of Animals: 2020 Edition* **30 BT-A**, 121 (2020).

55. Fleming, S. A., Monaikul, S., Mudd, A. T., Jacob, R. & Dilger, R. N. Extraction and dissection of the domesticated pig brain. *J. Vis. Exp.* (2021).

56. Lam, F. & Liang, Z.-P. A subspace approach to high-resolution spectroscopic imaging. *Magn. Reson. Med.* **71**, 1349–1357 (2014).

57. Nguyen, H. M., Peng, X., Do, M. N. & Liang, Z. Denoising MR spectroscopic imaging data with low-rank approximations. *IEEE Trans. Biomed. Eng.* **60**, 78–89 (2013).

Acknowledgements

The authors thank Dean Olson and Lingyang Zhu for discussions on ^{13}C NMR spectroscopy and acquiring the ^{13}C NMR data for the synthesized oligonucleotides; Furong Sun for developing the LC-MS/MS protocol and instruction on mass spectrometry data analysis; Yibo Zhao for his help with ^{13}C MRSI data processing; Adam Jones for his help on piglet feeding and rearing; Joanne Fil for her help on piglet brain tissue dissection; and Katie Ranard for her help on the DNA methylation measurement using ELISA kits. Thanks also to the Beckman Institute for Advanced Science and Technology and the Carl R. Woese Institute for Genomic Biology for access to facilities and technical support.

Author Contributions

FL contributed to concept development, technical implementation, data collection, data analysis, and manuscript writing. JC contributed to technical implementation, data collection, data analysis, and manuscript writing. JSC contributed to technical implementation, data collection, and data analysis. CC contributed to technical implementation, data collection, data analysis and manuscript writing. TKH contributed to data collection, data analysis, and manuscript writing. SKS contributed to concept development, technical implementation, data collection, data analysis, and manuscript

writing. ZPL contributed to concept development, technical implementation, data analysis, and manuscript writing. RND contributed to technical implementation, data collection, data analysis, and manuscript writing. GER contributed to concept development, technical implementation, data analysis, and manuscript writing. KCL contributed to concept development, technical implementation, data analysis, and manuscript writing.

Supplementary Information

Figure S1: LC-MS/MS data reveal ^{13}C-5mdC and 5mdC for quantifying labeling percentage. Representative time of flight (TOF) MS/MS multiple reaction monitor (MRM) data of **a**. ^{13}C-5-methylcytosine (fragmented from ^{13}C-5mdC after the first MS), **b.** 5-methylcytosine (fragmented from 5mdC), and **c**. cytosine (fragmented from dC). The data here were from genomic DNA isolated from the thalamus of a 10 days ^{13}C-labeled piglet. Characteristic elution times and specific molecular weights of the primary ions confirm the identities of the peaks. Area under the peaks were calculated to estimate the ^{13}C-enrichment of genomic 5mdC. Elution time are in minutes and peak area are indicated for each peak. 100% signal intensity is set to the intensity of the most prominent peak of the queried molecular weight.

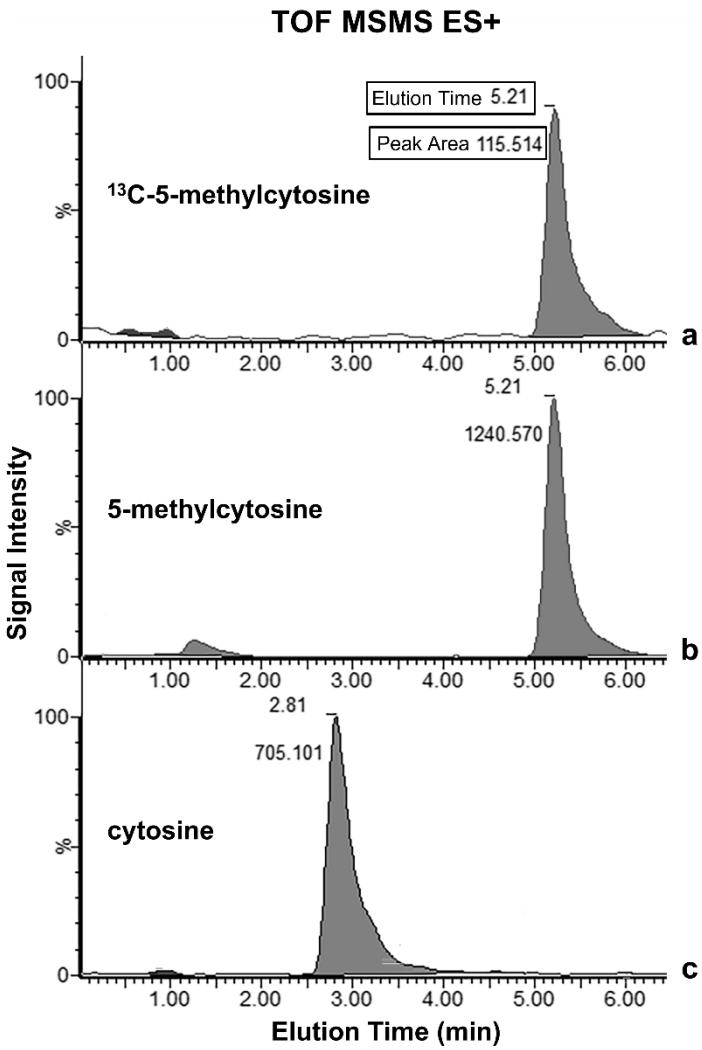

Figure S2: **^{13}C NMR spectroscopy specifically detects the ^{13}C signals from the 5mdC within a synthetic DNA oligonucleotide** (sequence: 5'-CTACGCCTCGCTCGCCCCTT-3'). **a**. High-resolution ^{13}C NMR spectrum of the unbuffered sample, acquired with a Bruker Carver B500 NMR spectrometer, 500 MHz (~11.75 T), equipped with a CryoProbe. The 10 slightly dispersed peaks at ~15 ppm from the methyl groups on the 5mdC nucleotides at different positions of the sequence are clearly identified and differentiated from the 5 peaks at ~14.2 ppm from the methyl groups on the dT nucleotides. **b**. The spectrum from the same sample buffered with HEPES to pH 7.4 showed consistent results, with the chemical shift of 5mdC at ~12.5 ppm. **c**. The spectrum from a control sequence for which all 5mdC were replaced with regular dC; all peaks from 5mdC disappeared, with only peaks from dT remaining. The black arrows in the first two rows indicate peaks from ethanol, which was removed in the last sample. These data provided the information that NMR signals for the ^{13}C-labeled 5mdC should be in the range of 12-15 ppm. On this basis, we assigned the increased signal observed at ~15 ppm in the brain experiments to ^{13}C-5mdC.

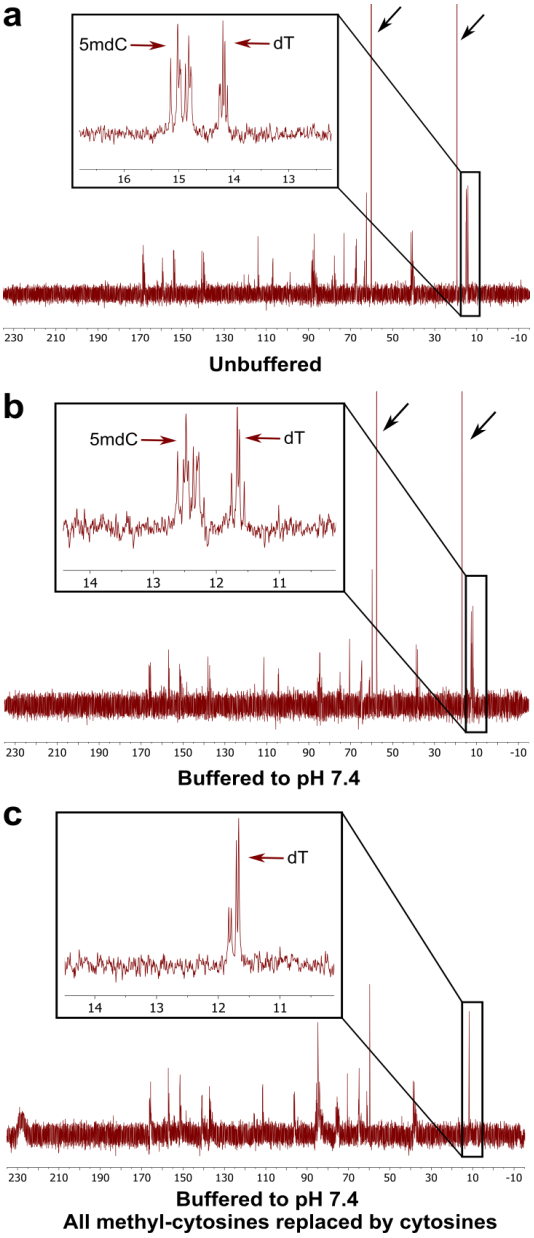

Figure S3: ^{13}C NMR spectroscopy data and fitting results from a 32-day labeled brain. **a**. High-fidelity fitting of the whole-sample SVS data. The blue curve is the real part of the raw ^{13}C NMR spectrum, and the black curve is the model fit. The fitting residual shown below in red is at noise level and passed the Komolgrov Gaussianity test. **b**. Different spectral components extracted from the fitting (coded by different colors and with lineshape distortion $h(t)$ incorporated). These components were used to construct the basis for fitting the MRSI data. The blue peak (right column) corresponds to ^{13}C-5mdC from labeling. The strong green peak centered at ~55 ppm is hypothesized to be from ^{13}C labeled phospholipids through the methylation pathway that converts phosphatidylethanolamine to phosphatidylcholine. This signal is not of interest in our eMRI study and was thus excluded in Fig. 3. But it may be of interest in studies concerning lipid metabolism.

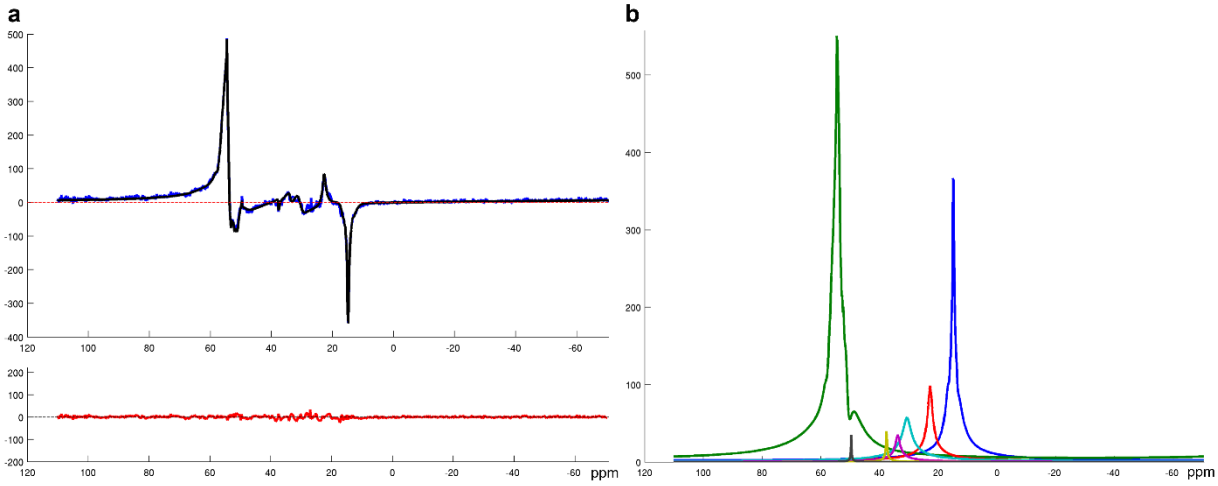

Figure S4: ^{13}C-5mdC maps across the whole 3D imaging volume. The spatiospectrally reconstructed ^{13}C-MRSI data from Fig. 4 were first resized to the resolution of the anatomical MPRAGE image for better visualization and then quantified. The quantified ^{13}C-5mdC maps are shown in gray scale, normalized to [0,1]. Strong signals can be seen within the brain, with a clear difference from the background noise.

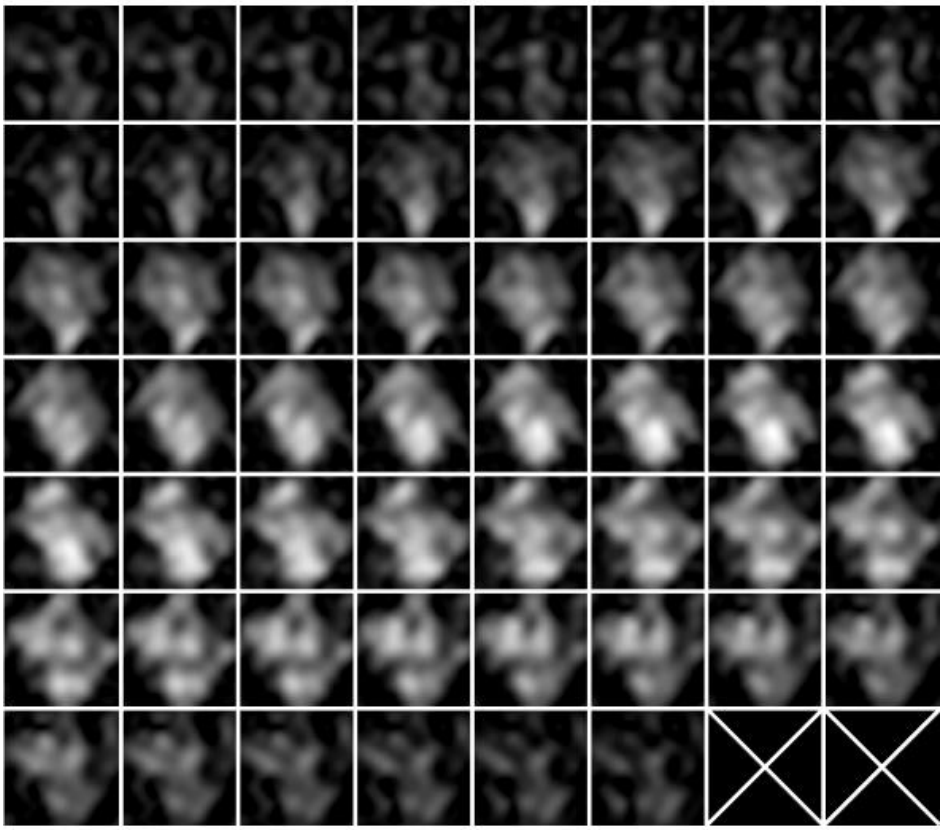

Figure S5: Single repetition ^{13}C-MRSI versus five-repetition ^{13}C-MRSI (5× longer imaging time). Each repetition in our ^{13}C imaging scan took about 18 h. While the mapping results in Fig. 4 were from a five-repetition scan, as shown by the image below, the regional eMRI measurements from a single repetition are consistent with a 5× longer scan (correlation coefficient ~0.9). This indicates strong translational potential to human experiments. Taking into account the larger brain volume for human, lower labeling efficiency, and the use of a 7T human system, we predict an approximately 1-2 h scan would be necessary for high-quality human data. Meanwhile, it is important to note that since eMRI measures a very stable signal due to the stable ^{13}C isotope in the DNA, we have the flexibility to acquire multiple signal averages across several shorter experiment sessions to achieve a sufficient SNR.

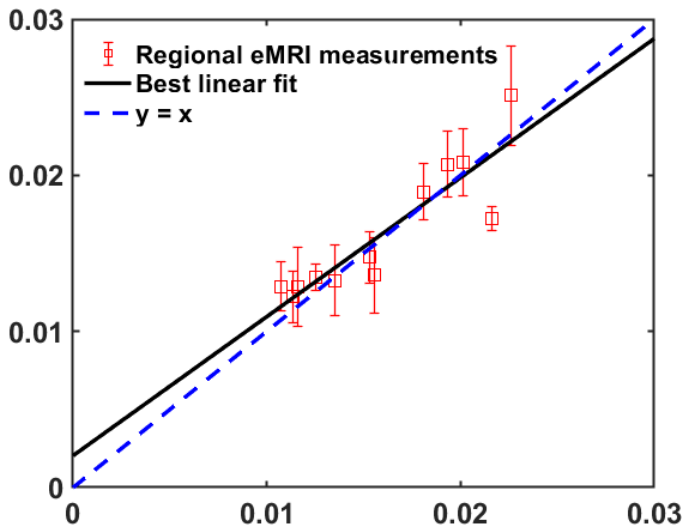

Figure S6: Phantom results for validating our ^{13}C-MRSI acquisition and processing methods. **a**. Phantom setup: Four 10 mm NMR tubes were filled with ^{13}C-5mdC solutions with varying concentrations and placed in the microimaging system. The concentrations are labeled in the T2-weighted MRI shown. **b**. A ^{13}C-5mdC map for one 2D slice from the 3D imaging data, with a spatial resolution of 3.75 × 3.75 × 5 mm^3. The spatial intensity distribution matches well with the true concentration variations. **c**. Quantitative comparison between regional concentration estimates from individual tubes and the true values. The original estimates were in arbitrary units. The highest value was normalized to 10 mM and regressed against the true values. Accurate results were obtained with a 0.9 correlation coefficient between the estimates and true values. The error bars of the measurements indicate the standard deviations within each region of interest.

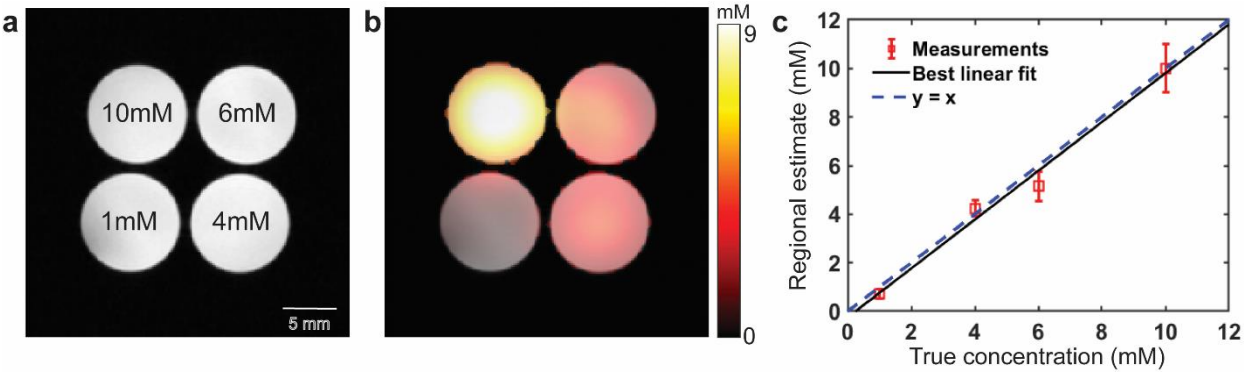

Figure S7: Global DNA methylation in the putamen (Put) and temporal lobe (TL) measured by ELISA is consistent with eMRI data. Relative methylation levels in the Put and TL were measured in two different trials using a fluorometric ELISA kit. Since the goal is to confirm regional difference and the absolute methylation levels are not of interest, we calculated the ratios of methylation levels between Put and TL, with TL methylation level normalized to 1. Both trials showed clearly higher methylation in Put than TL with ratios > 1, concurring with the regional relationship measured in eMRI. The error bars indicate the standard deviations within each brain region. While higher Put methylation was consistently measured, the exact ratios varied between trials 1 to 2 (run in consecutive days) due to fluorescence reading variations. More reproducible kits and protocols will be evaluated in follow-up research when comparing global methylation across more regions.

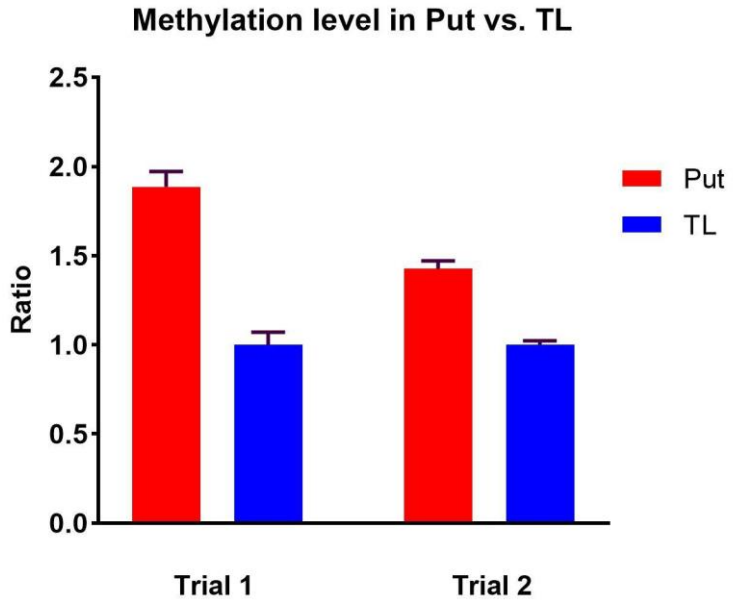
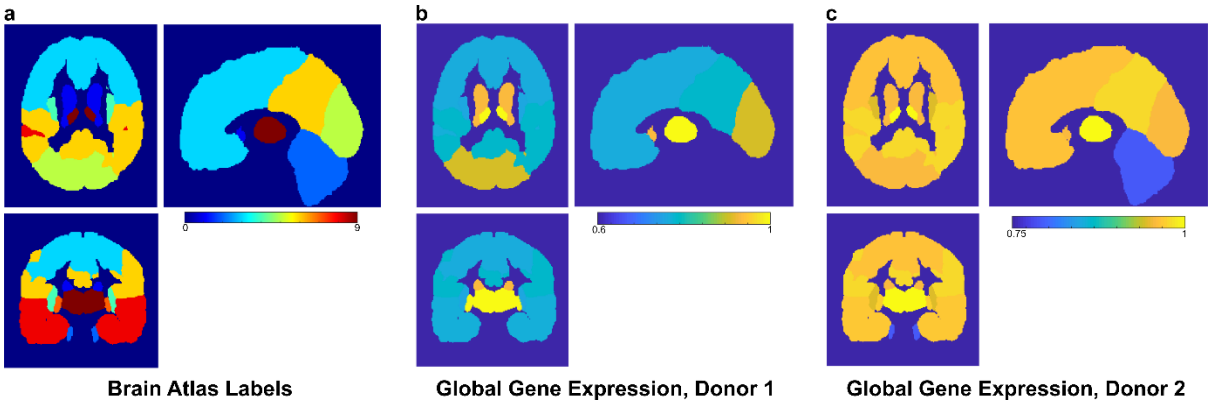
Figure S8. Analysis of the Allen Brain Atlas (ABA) reveals spatial variation in global gene expression in human brains. RNA-sequencing (RNAseq) data from two brains were downloaded from the ABA (https://human.brain-map.org/static/download). The RNASeq data contain transcriptomic profiles for tissues sampled from different brain regions and were used to calculate global gene expression. These global expression values were then mapped to a brain atlas (http://nist.mni.mcgill.ca/mni-average-brain-305-mri/). **a**. A three-plane view of the MNI brain atlas was shown, with different brain regions segmented from MRI and labeled. Nine anatomical regions were considered: caudate, cerebellum, frontal lobe, insula, occipital lobe, parietal lobe, putamen, temporal lobe, and thalamus. **b**. Maps of global gene expression from one of the two ABA brains. The RNAseq data contain expression data for 22318 genes, each of which has measurements for tissues sampled across different brain structures. We kept only the genes for which the expression levels were reliably detected in all samples. For each brain structure, global gene expression was calculated by summing the expression levels of individual genes, which provided an array of values that were then mapped to the MNI atlas regions in a. Clear regional variation in global gene expression can be visualized. **c**. The same analysis is shown for the second ABA brain. The two brains were from donors at different ages. The global expression values across different structures were normalized using the maximum values for both b and c. Only the brain structures available in both the MNI atlas and RNASeq data were considered. These results support the presence of spatial variation in global gene expression and underline the importance of an *in vivo* surrogate (eMRI) to explore the functional significance of this variation for brain function and disease studies.