PNAS
www.pnas.org

Supplementary Information for

Epigenetic MRI: Noninvasive Imaging of DNA Methylation in the Brain

Fan Lam, James Chu, Ji Sun Choi, Chang Cao, T. Kevin Hitchens, Scott K. Silverman, Zhi-Pei Liang, Ryan N. Dilger, Gene E. Robinson* and King C. Li*

*Gene E. Robinson and King C. Li

Email: generobi@illinois.edu and kingli@illinois.edu

This PDF file includes:

Figures S1 to S8

Supplementary Figures

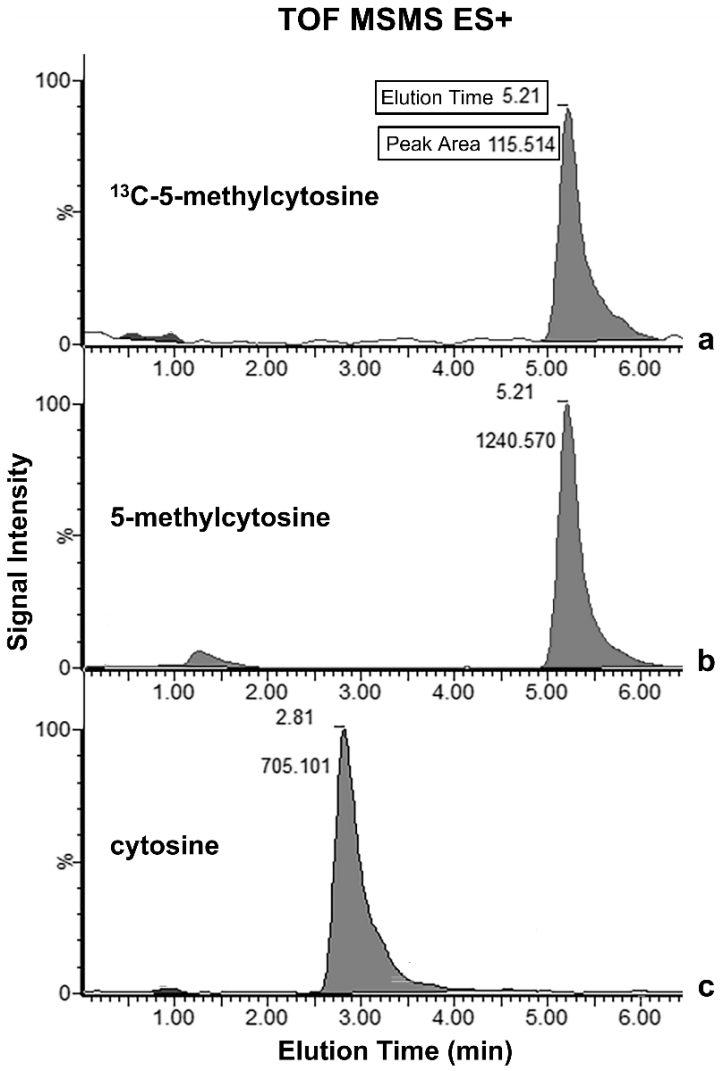

Fig. S1. LC-MS/MS data reveal ^{13}C-5mdC and 5mdC for quantifying labeling percentage. Representative time of flight (TOF) MS/MS multiple reaction monitor (MRM) data of **a**. ^{13}C-5-methylcytosine (fragmented from ^{13}C-5mdC after the first MS), **b.** 5-methylcytosine (fragmented from 5mdC), and **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.

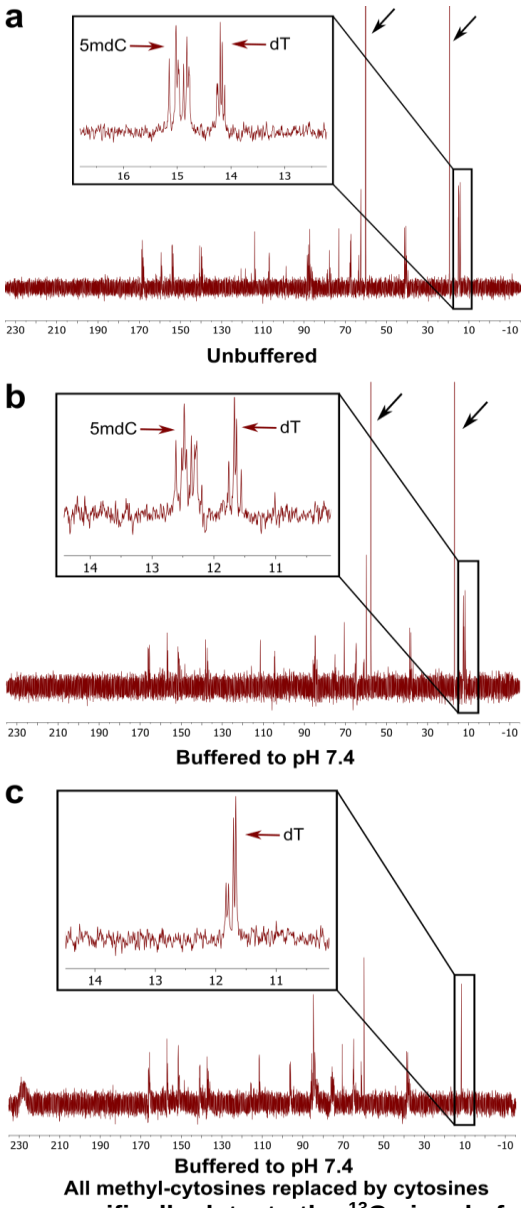

Fig. S2. ^{13}C NMR spectroscopy specifically detects the ^{13}C signals from the 5mdC within a synthetic DNA oligonucleotide (sequence: 5'-CTACGCCTCGCTCGCCCCTT-3'). **a**. High-resolution ^{13}C NMR spectrum of the unbuffered sample, acquired with a Bruker Carver B500 NMR spectrometer, 500 MHz (~11.75 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.

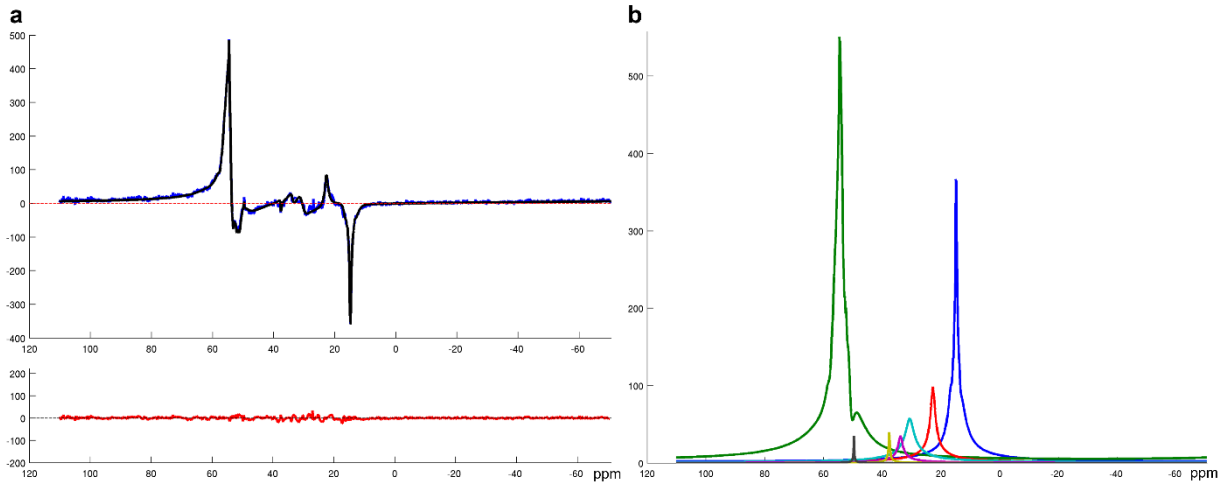

Fig. S3. ^{13}C NMR spectroscopy data and fitting results from a 32-day labeled brain. **a**. High-fidelity fitting of the whole-sample SVS data. The blue curve is the real part of the raw ^{13}C NMR spectrum, and the black curve is the model fit. The fitting residual shown below in red is at noise level 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 the present study and was thus excluded in Fig. 3. But it may be of interest in studies concerning lipid metabolism.

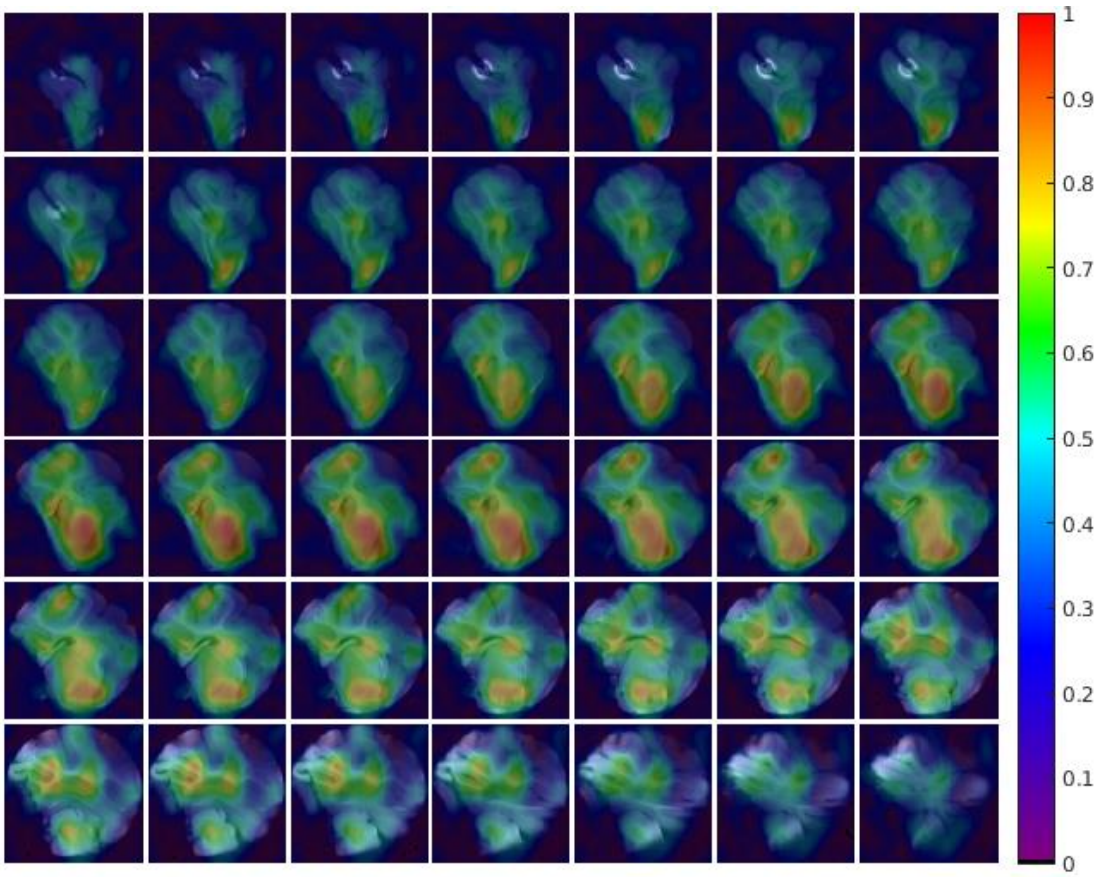

Fig. S4. ^{13}C-5mdC maps across the whole 3D imaging volume. The spatiospectrally reconstructed ^{13}C-MRSI data (with a lower resolution than the MR images) from Fig. 4 were first resized to the resolution of the anatomical MR images and then quantified (for better visualization). The quantified ^{13}C-5mdC maps were color coded and overlaid on the 3D MRI, normalized to [0,1]. Strong signals can be seen within the brain, with a clear distinction from the nonbrain background noise, implying an excellent signal-to-noise ratio.

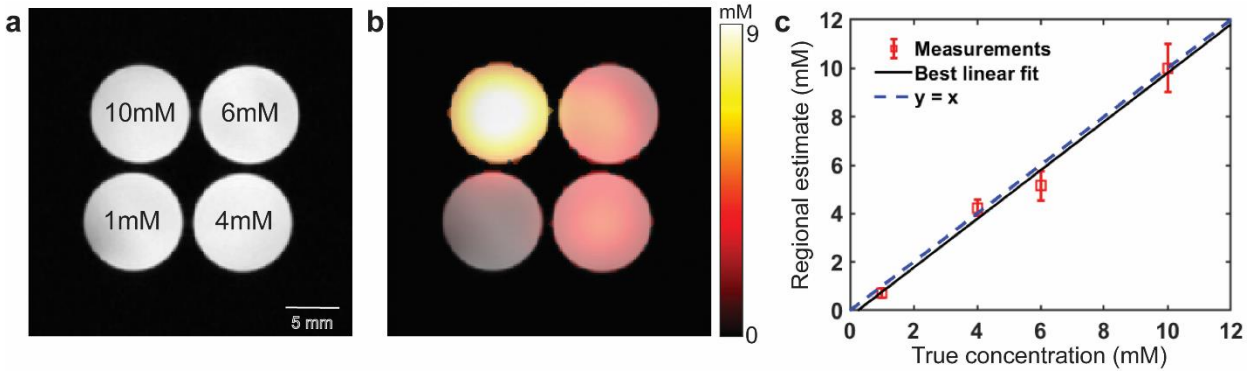

Fig. S5. Phantom results for validating our ^{13}C-MRSI acquisition and processing methods. **a**. Phantom setup: Four 10 mm NMR tubes were filled with ^{13}C-5mdC solutions with varying concentrations and placed in the microimaging system. The concentrations are labeled in the T2-weighted MRI shown. **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.

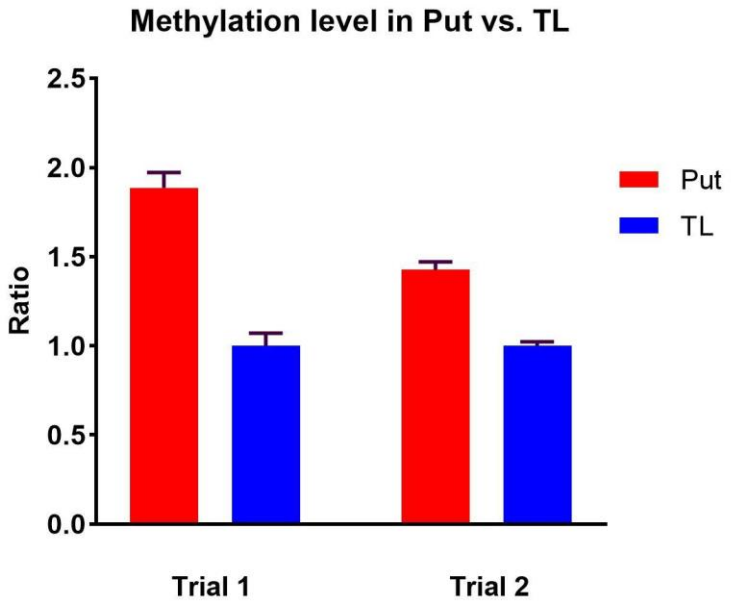

Fig. S6. Global DNA methylation in the putamen (Put) and temporal lobe (TL) measured by ELISA is consistent with eMRI data. Relative methylation levels in the Put and TL were measured in two different trials using a fluorometric ELISA kit. Since the goal was to confirm regional differences and the absolute methylation levels were therefore 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, consistent with the results from eMRI. 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.

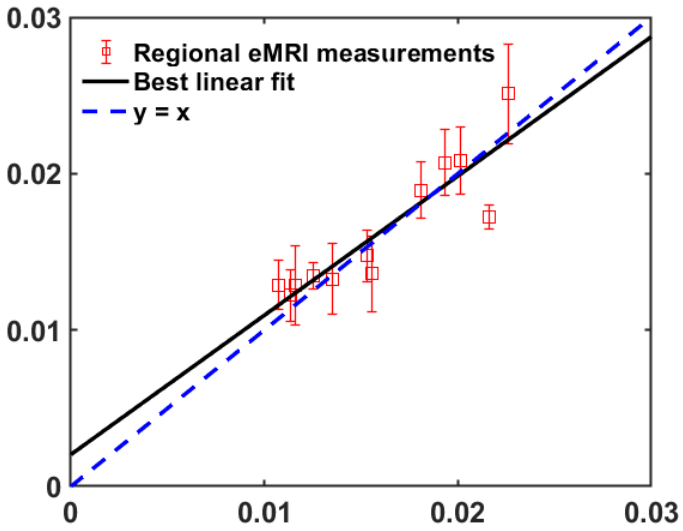

Fig. S7. Single repetition ^{13}C-MRSI versus five-repetition ^{13}C-MRSI (5× longer imaging time). Each repetition in our ^{13}C imaging scan took about 18 h. While the mapping results in Fig. 4 were from a five-repetition scan, as shown by the image below, the regional eMRI measurements from a single 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 7 T human system, we predict an approximately 1-2 h scan would be sufficient for high-quality human data. However, since eMRI measures a very stable signal due to the stability of the ^{13}C isotope, we have the flexibility to acquire multiple signal averages across several shorter experiment sessions to achieve a sufficient signal-to-noise ratio.

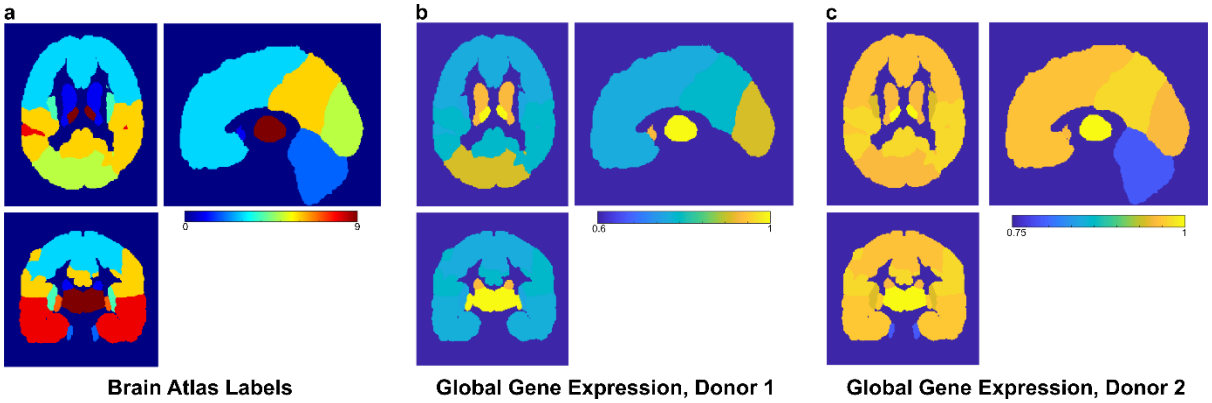

Fig. S8. Analysis of the Allen Brain Atlas (ABA) reveals spatial variation in global gene expression in human brains. RNA-sequencing (RNAseq) data from two brains were downloaded from the ABA (https://human.brain-map.org/static/download). The RNASeq data contain transcriptomic profiles for tissues sampled 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 Montreal Neurological Institute (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, consistent with our eMRI results. eMRI provides a non-invasive surrogate to explore the functional significance of variation in gene expression for brain function and disease.