Mice carrying a human *GLUD2* gene recapitulate aspects of human transcriptome and metabolome development

Qian Li[a,b,1], Song Guo[a,1], Xi Jiang[a], Jaroslaw Bryk[c,2], Ronald Naumann[d], Wolfgang Enard[c,3], Masaru Tomita[e], Masahiro Sugimoto[e], Philipp Khaitovich[a,c,f,4], and Svante Pääbo[c,4]

[a]Chinese Academy of Sciences Key Laboratory of Computational Biology, Chinese Academy of Sciences-Max Planck Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 200031 Shanghai, China; [b]University of Chinese Academy of Sciences, 100049 Beijing, China; [c]Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany; [d]Max Planck Institute of Molecular Cell Biology and Genetics, D-01307 Dresden, Germany; [e]Institute for Advanced Biosciences, Keio University, 997-0035 Tsuruoka, Yamagata, Japan; and [f]Skolkovo Institute for Science and Technology, 143025 Skolkovo, Russia

Edited by Joshua M. Akey, University of Washington, Seattle, WA, and accepted by the Editorial Board April 1, 2016 (received for review September 28, 2015)

Whereas all mammals have one glutamate dehydrogenase gene (*GLUD1*), humans and apes carry an additional gene (*GLUD2*), which encodes an enzyme with distinct biochemical properties. We inserted a bacterial artificial chromosome containing the human *GLUD2* gene into mice and analyzed the resulting changes in the transcriptome and metabolome during postnatal brain development. Effects were most pronounced early postnatally, and predominantly genes involved in neuronal development were affected. Remarkably, the effects in the transgenic mice partially parallel the transcriptome and metabolome differences seen between humans and macaques analyzed. Notably, the introduction of *GLUD2* did not affect glutamate levels in mice, consistent with observations in the primates. Instead, the metabolic effects of *GLUD2* center on the tricarboxylic acid cycle, suggesting that *GLUD2* affects carbon flux during early brain development, possibly supporting lipid biosynthesis.

human evolution | *GLUD2* | brain metabolism

Glutamate dehydrogenase (GDH) is a metabolic enzyme catalyzing the conversion of glutamate to α-ketoglutarate and ammonia (1). Whereas the ammonia is metabolized via the urea cycle, the α-ketoglutarate enters the tricarboxylic acid (TCA) cycle in mitochondria. Besides its metabolic role, glutamate also functions as a major excitatory neurotransmitter (2, 3).

Whereas most organisms contain one copy of the *GLUD* gene encoding the GDH enzyme, humans and apes have two: *GLUD1* and an additional gene, *GLUD2*, which originated by retroposition of the *GLUD1* transcript after the split from apes and old world monkeys (4). Sequence analysis suggests that it is highly unlikely that *GLUD2* would have contained an intact ORF and a K_a/K_s <1 throughout the evolution of apes without being functional (5). Moreover, positive selection has affected *GLUD2* (4), including changes in amino acid residues that make it less sensitive to low pH and GTP inhibition, and resulting in a requirement for high ADP levels for allosteric activation (4, 6, 7). *GLUD2* mRNA expression levels in tissues are lower than those of *GLUD1*, but similarly distributed across tissues. However, whereas the ancestral version of the *GLUD* enzyme occurs both in mitochondria and the cytoplasm, *GLUD2* is specifically targeted to mitochondria (8, 9).

Changes in *GLUD2* properties have been suggested to reflect functional adaptation to the metabolism of the neurotransmitter glutamate in the brain (4, 10), and the fact that *GLUD2* has been positively selected and maintained during ape and human evolution suggests that it may have physiological effects important for the function of ape and human brains. However, the connection between the emergence of the *GLUD2* gene in the ancestors of apes and humans and changes in brain function remains elusive. To date, the only direct insights into *GLUD2* function come from a rare *GLUD2* mutation linked to the onset of Parkinson's disease (11) and from glioma cells carrying a mutated isocitrate dehydrogenase 1 gene (*IDH1*) where *GLUD2* expression reverses the effects of the *IDH1* mutation by reactivation of the metabolic flux from glucose and glutamine to lipids by way of the TCA cycle (12).

To investigate the physiological role the *GLUD2* gene may play in human and ape brains, we generated mice transgenic for a genomic region containing human *GLUD2*. We compared effects on gene expression and metabolism during postnatal development of the frontal cortex of the brain in these mice and their wild-type littermates, with similar data obtained from humans and rhesus macaques, the closest relative of humans and apes, which lack the *GLUD2* gene.

Results

A Mouse Model of *GLUD2*. We constructed transgenic mice carrying a 176-kb-long human genomic region containing the *GLUD2* gene, as well as 43 kb of upstream and 131 kb of downstream DNA sequences (*SI Appendix*, Fig. S1). To account for effects potentially caused by stochastic insertion of *GLUD2* sequence into the mouse

Significance

A novel version of the glutamate dehydrogenase gene, *GLUD2*, evolved in the common ancestors of humans and apes. Based on sequence and expression pattern, *GLUD2* has been suggested to play a role in glutamate metabolism in human and ape brains. We have generated transgenic mice carrying a human *GLUD2* gene. Analysis of transcriptome and metabolome changes induced by *GLUD2* in the cerebral cortex revealed no changes in glutamate concentration but instead changes to metabolic pathways centering on the TCA cycle during early postnatal development. These changes mirrored differences seen between human and macaque during cortex development, suggesting that *GLUD2* may play a role during brain development in apes and humans, possibly by providing precursors for the biosynthesis of lipids.

Author contributions: W.E., P.K., and S.P. designed research; X.J., J.B., R.N., and M.S. performed research; W.E., M.T., and M.S. contributed reagents/analytic tools; Q.L. and S.G. analyzed data; and Q.L., P.K., and S.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. J.M.A. is a guest editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE80122).

[1]Q.L. and S.G. contributed equally to this work.

[2]Present address: School of Applied Sciences, University of Huddersfield, HD1 3DH Huddersfield, UK.

[3]Present address: Department of Biology II, Ludwig Maximilians University, 82152 Martinsried, Germany.

[4]To whom correspondence may be addressed. Email: khaitovich@eva.mpg.de or paabo@eva.mpg.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1519261113/-/DCSupplemental.

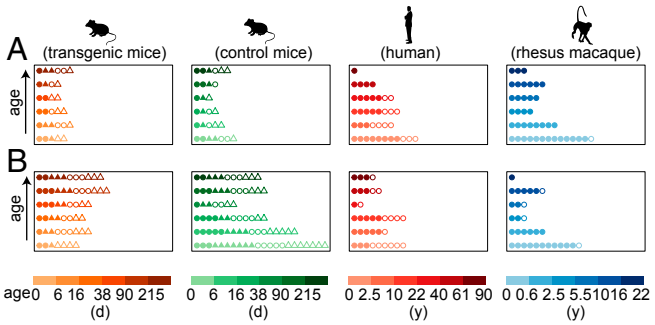

Fig. 1. Sample information. (*A*) Schematic representation of age, sex, and transgenic line information for mouse and primate samples used for transcriptome (RNA-seq) measurements. (*B*) Schematic representation of age, sex, and transgenic line information for mouse and primate samples used for metabolome (CE-MS) measurements. Each symbol represents an individual sample. The colors indicate genotype/species information: orange, transgenic mice; green, control mice; red, humans; blue, rhesus macaques. Lighter shades of color indicate younger ages. The bars show the corresponding age intervals: d, days; y, years. The symbols indicate mice lines: circles, line a; triangles, line b. Filled symbols indicate males; empty symbols indicate females.

genome, we constructed two independent transgenic lines (a and b). Anatomical, neurophysiological, and behavioral analyses of adult mice did not reveal any overt effects of the *GLUD2* genotype in the two lines (13).

We assessed the effects of *GLUD2* on gene expression and metabolite concentrations in the developing frontal cortex of hemizygous ($GLUD2^+$) mice from both lines as well as control littermates, using RNA sequencing (RNA-seq) and capillary electrophoresis coupled with mass spectrometry (CE-MS). RNA-seq data were collected from 30 $GLUD2^+$ and 29 control individuals. Metabolite concentrations were measured in 56 $GLUD2^+$ and 81 control individuals. The mice varied in age from 3 d to 18 mo, with ages and sexes matched between transgenic and control groups (Fig. 1 *A* and *B* and *SI Appendix*, Table S1).

To assess whether molecular changes induced by *GLUD2* in the mice might recapitulate differences between hominoids (apes and humans) and other primates, we measured metabolite concentrations in prefrontal cortex of 35 humans and 31 rhesus macaques using CE-MS. For both species, samples covered the respective lifespans approximately corresponding to the one sampled in the mice: from 16 d postnatal to 90 y in humans, and from 18.4 wk postconception to 21 y in macaques (Fig. 1*B* and *SI Appendix*, Table S2). We furthermore analyzed RNA-seq data previously collected in the prefrontal cortex of 38 humans and 40 rhesus macaques covering the same age span (14) (Fig. 1*A* and *SI Appendix*, Table S2). Between these two datasets, 13 human and 20 macaque individuals overlapped for both the RNA-seq and CE-MS data (*SI Appendix*, Table S2).

***GLUD2* Affects Gene Expression During Early Postnatal Development.** Based on 613 million, 100 nt-long sequence reads collected from the mouse polyA-plus RNA (*SI Appendix*, Table S3), we detected expression of 18,610 protein-coding and noncoding genes. Global expression variation among the samples showed a clear effect of age when analyzed by principal component analysis (PCA), whereas neither genotype nor sex affected gene expression substantially (Fig. 2*A* and *SI Appendix*, Fig. S2). Consistently, as many as 15,663 of the 18,610 detected genes, including the *GLUD2* transgene, showed significant age-dependent expression changes in the mouse cortex [*F*-test, Benjamini–Hochberg (BH) corrected $P < 0.05$, false discovery rate (FDR) < 0.01].

Both transgenic mouse lines showed the same developmental expression profile of *GLUD2*, which was distinct from the expression profile of *GLUD1* (*SI Appendix*, Fig. S3). Notably, the *GLUD2* expression trajectory in the developing mouse brain strongly resembled the trajectory of *GLUD2* expression during prefrontal cortex development in humans (Pearson correlation, $r = 0.86$, $P < 0.0001$; *SI Appendix*, Fig. S4). We also detected expression of a long noncoding RNA (lncRNA) originating from the opposite strand of the transfected human genomic region upstream of the *GLUD2* transcription start site (*SI Appendix*, Fig. S5). Expression profile of this lncRNA closely resembled the expression profile of *GLUD2* (*SI Appendix*, Fig. S6), suggesting that it is expressed from the same bidirectional promoter as *GLUD2* (*SI Appendix*, Table S4).

An effect of the *GLUD2* genotype, albeit small, was detectable: using analysis of covariance (ANCOVA) with linear, quadratic, and cubic models, we identified 12 protein-coding genes and one lncRNA showing significant expression differences between the two transgenic lines and control littermates (*F*-test, BH corrected $P < 0.05$, permutations $P < 0.05$, FDR < 0.01, *SI Appendix*, Table S5). Strikingly, the effect of the *GLUD2* genotype was observed only during early postnatal development and disappeared at approximately 2 wk of age (Fig. 2*B*). By contrast, other genes expressed in the mouse brain did not show any increased divergence between transgenic and control mice during early development (Fig. 2*C*).

Eleven of the 13 genes differentially expressed between transgenic and control mice fell into the same coexpressed module in an unsupervised hierarchical clustering analysis (Fig. 3*A* and *SI Appendix*, Table S5 and Fig. S7), an observation not expected by chance (permutations, $P < 0.001$) (*SI Appendix*, Fig. S8). The genes affected by *GLUD2* are thus coexpressed during mouse development. The expression of these 11 genes decreased rapidly during early postnatal development in both transgenic and control mice, but in the transgenic mice, the decrease in expression was shifted to earlier stages of development (Fig. 3 *B* and *C* and *SI Appendix*, Fig. S9).

Based on analysis of Gene Ontology (GO) terms (15), the 11 coexpressed genes were significantly enriched in several biological processes, many of them related to neural development and transcriptional regulation (hypergeometric test, FDR corrected $P < 0.05$) (Fig. 3*D* and *SI Appendix*, Table S6), a result robust to the use of different background gene distributions. The genes affected by *GLUD2* may thus to some extent be functionally related.

***GLUD2* Effects in the Mice Mirror Differences Between Primates.** To assess whether the effect of *GLUD2* in the mice recapitulated differences between primate species that lack and that have *GLUD2*, we reanalyzed 579 million, 100-nt-long RNA-seq reads from the prefrontal cortex of rhesus macaques and humans. We detected the expression of 23,115 protein-coding and noncoding

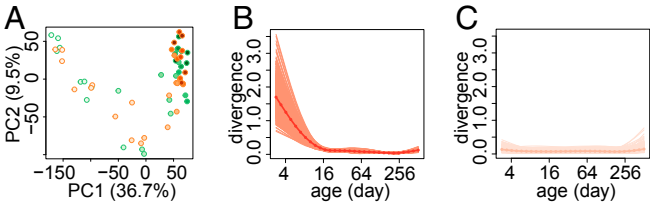

Fig. 2. Effect of the *GLUD2* genotype on gene expression in the mouse model. (*A*) The first two principal components (PCs) of the principal component analysis, based on the expression of 18,610 genes detected in the mouse brain. Each circle represents an individual. The colors indicate genotype information: orange, transgenic mice; green, control mice. Symbols filled by lighter shades of color indicate younger ages. (*B*) The mean normalized gene expression divergence between transgenic and control mice, based on the 13 age-dependent genes with expression affected by the *GLUD2* genotype (red curve). The colored area shows variation of the divergence estimates obtained by bootstrapping the 13 genes 1,000 times. (*C*) The mean normalized gene expression divergence between transgenic and control mice, based on the remaining 15,650 age-dependent genes (light-red curve). The colored area shows variation of the divergence estimates obtained by subsampling 13 genes 1,000 times.

EVOLUTION

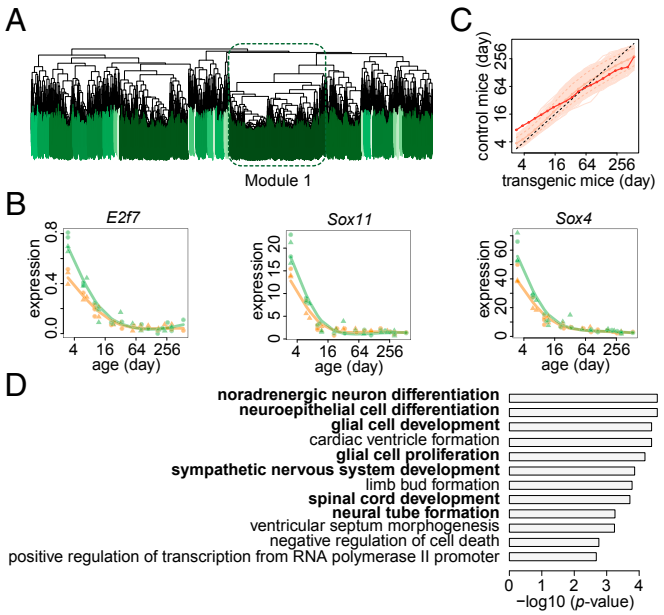

Fig. 3. Patterns and functions of genes with expression affected by *GLUD2* genotype. (*A*) Dendrogram based on unsupervised hierarchical clustering of 15,663 age-dependent genes. The colors represent coexpressed modules. The darker shades of color indicate larger modules. Dashed green rectangle shows module 1 containing 11 of the 13 genes with expression affected by the *GLUD2* genotype. (*B*) Expression profiles of 3 of the 11 module 1 genes with expression affected by the *GLUD2* genotype. Each point represents an individual sample. The colors indicate genotype information: orange, transgenic mice; green, control mice. The symbols indicate mice lines: circles, line a; triangles, line b. The lines show spline curves fitted to expression data points with four degrees of freedom. (*C*) Developmental time shift between transgenic mice and control mice, calculated based on expression of 11 module 1 genes with expression affected by the *GLUD2* genotype (dark-red curve) and the remaining 15,652 age-dependent genes (light-red curve). The light-red–colored area shows variation of time-shift estimates for the remaining age-dependent genes, obtained by randomly subsampling 11 genes 1,000 times. The curves are obtained by aligning transgenic mice expression time series to that of control mice, using the modified dynamic time warping algorithm, showing the ages where transgenic mice expression levels correspond to those of control mice. During early stages of development, transgenic mice ages are mapped to older ages in control mice, indicating transgenic mice expression profiles shift to earlier developmental stages. (*D*) GO biological processes significantly enriched in 11 module-1 genes with expression affected by the *GLUD2* genotype. Biological processes related to neural development are shown in bold.

genes. PCA analysis of the expression of these genes demonstrated substantial species-dependent as well as age-dependent divergence (Fig. 4*A*).

Among the 13 genes differentially expressed between transgenic and control mice, 9 were expressed in human and macaque prefrontal cortex (*SI Appendix*, Table S5). Remarkably, the expression of these 9 genes showed a large divergence between humans and macaques at the earliest stages of postnatal development after which the divergence rapidly decreased. Thus, the expression differences between humans and macaques resembled those seen between mice carrying a *GLUD2* gene and control littermates (Fig. 4*B*). By contrast, other genes expressed in human and macaque prefrontal cortex showed no such divergence pattern (permutation $P = 0.05$, Fig. 4*C*).

Furthermore, seven of the nine genes showed a rapid expression level decrease during human and macaque early postnatal development, i.e., the same developmental expression trajectory as the genes affected by *GLUD2* in the transgenic mice (Pearson correlation, $r > 0.9$, permutation $P < 0.001$, *SI Appendix*, Figs. S10 and S11). After adjusting for the differences in lifespan between humans and macaques, we found a shift in the timing of the expression to earlier developmental stages in humans relative to macaques (Fig. 4*D* and *SI Appendix*, Fig. S10), similar to what is seen in the *GLUD2* mice. This was not caused by any general mismatching of human and macaque developmental timing, as other genes did not show such a coordinated shift to earlier developmental stages (Pearson correlation, $r > 0.9$, permutation $P = 0.02$, *SI Appendix*, Fig. S12). Thus, gene expression changes in *GLUD2* mice indeed recapitulate gene expression differences between human and macaque expression profiles during early postnatal development.

Among the 7 genes showing similar expression and developmental shifts in the transgenic mice and in human cortex, 3 are transcription factors (TFs) implicated in neural development: SRY (sex determining region Y)-box 4 (*SOX4*), SRY (sex determining region Y)-box 11 (*SOX11*), and Distal-less homeobox 2 (*DLX2*) (16–22). A total of 15 known neural-related target genes of these TFs (16–29) were expressed in the mice and 14 in the primates (*SI Appendix*, Table S7). The expression divergence profiles of these target genes and the corresponding TFs between the transgenic and control mice and between humans and macaques at different ages were significantly correlated (Pearson correlation, $r > 0.5$; permutation $P = 0.008$ for mice and $P = 0.001$ for primates, *SI Appendix*, Fig. S13), suggesting that the effects of *GLUD2* may go beyond the 13 genes showing differences between the transgenic and control mice.

As the effects of *GLUD2* in the mice were strongest at the earliest stages of postnatal development, it is possible that *GLUD2* plays a role mainly in prenatal development. To test if this may be the case in humans, we analyzed the expression of the 13 genes affected by *GLUD2* expression in the mice using RNA-seq data from prenatal prefrontal cortex development of humans (30) and macaques. Of the 13 genes, 9 had detectable expression in our human and macaque data, as well as in public fetal human cortex data. Strikingly, expression of the 9 genes was substantially more divergent between humans and macaques before than after birth (Fig. 5 *A* and *B* and *SI Appendix*, Fig. S14), whereas other

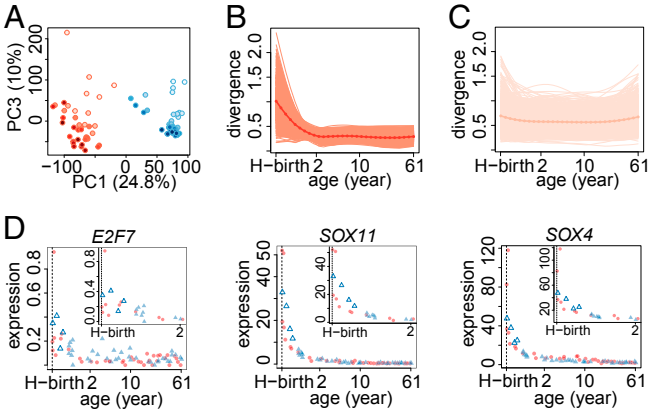

Fig. 4. Gene expression in primates. (*A*) The first and third PCs of the principal component analysis, based on the expression of 23,115 genes detected in human and rhesus macaque prefrontal cortex. Each circle represents an individual (red, humans; blue, macaques). Lighter shades of color indicate younger ages. (*B*) The mean normalized gene expression divergence between humans and macaques based on the 9 genes with expression affected by *GLUD2* genotype in the mouse model (red curve). The colored area shows variation of the divergence estimates obtained by bootstrapping the 9 genes 1,000 times. "H-birth" *x*-axis marks represent human birth age. (*C*) The mean normalized gene expression divergence between humans and macaques based on the remaining 23,106 detected genes (light-red curve). The colored area shows variation of the divergence estimates obtained by subsampling 9 genes 1,000 times. (*D*) Expression profiles of E2F transcription factor 7 (*E2F7*), *SOX11*, and *SOX4*, which are among the 7 genes showing the similar expression and divergence pattern in primates and the transgenic mice. Each point represents an individual. The colors indicate species information: red, humans; blue, macaques. Filled symbols indicate postnatal ages; empty points, prenatal ages. The vertical dashed line shows the human birth age.

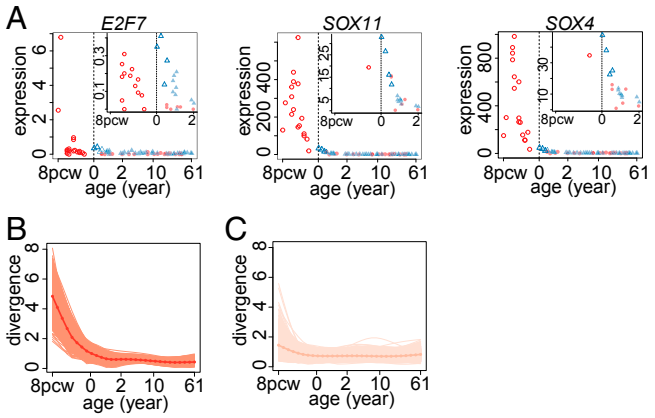

Fig. 5. Prenatal gene expression. (*A*) Prenatal expression profiles of *E2F7*, *SOX11*, and *SOX4* genes, which are among the 7 genes showing similar expression and divergence patterns in primates and the transgenic mice. Each point represents an individual. The colors indicate species information: red, humans; blue, macaques. Filled symbols indicate postnatal ages; empty points, prenatal ages (17 in humans and 4 in macaques). The vertical dashed line shows the human birth age; pcw, weeks postconception. (*B*) The mean normalized gene expression divergence between humans and macaques based on the 9 genes with expression affected by the *GLUD2* genotype, calculated using a published fetal human dataset (red curve). The colored area shows variation of the divergence estimates obtained by bootstrapping the 9 genes 1,000 times. (*C*) The mean normalized gene expression divergence between humans and macaques, based on the remaining 15,897 detected genes in the public fetal human and our macaque time series data (light-red curve). The colored area shows variation of the divergence estimates obtained by subsampling 9 genes 1,000 times.

expressed genes showed no obvious increase in expression differences before birth (Fig. 5*C*).

Brain Metabolism in *GLUD2* Mice and Primates. We assessed the effects of *GLUD2* on metabolism in the brains of the *GLUD2* mice using CE-MS. For comparisons, we similarly analyzed human and macaque brains. We detected and quantified 110 and 88 metabolites in the mice and in humans and macaques, respectively (*SI Appendix*, Tables S8 and S9). PCA analysis based on concentration levels of these metabolites revealed a substantial effect of age in both mice and primates (Fig. 6 *A* and *B*).

We first focused on the concentration levels of glutamate, the direct substrate of the GDH2 enzyme encoded by *GLUD2*. We detected no effect of *GLUD2* on glutamate concentration in the mouse cortex (Fig. 6*C*). By contrast, glutamate concentrations in the prefrontal cortex of the humans are substantially lower than in the macaques (Fig. 6*C*). Previous work using gas chromatography coupled with mass spectrometry (GC-MS) has shown that compared with humans, chimpanzees as well as macaques have higher glutamate levels in the brain (31), something that reexamination of the published glutamate data confirms (Fig. 6*C*). Because both humans and chimpanzees carry *GLUD2* genes, a mechanism other than the mere presence of *GLUD2* must be responsible for the lower glutamate concentrations in the human brain.

Despite the absence of a *GLUD2* effect on glutamate levels in the mice, we detected that the overall differences between the metabolomes of the transgenic and the control mice were about threefold larger shortly after birth relative to 2 mo after birth (*SI Appendix*, Fig. S15). To test if this may be at least partially connected to gene expression differences, we compared the concentration profiles of 24 metabolites associated with the metabolic pathways where the 13 genes differentially expressed between transgenic and control mice are located (*SI Appendix*, Table S8). Concentration differences of the 24 metabolites were substantially larger than the differences for the 86 remaining compounds analyzed (Fig. 7*A*). Similar results were obtained using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotation and the Small Molecule Pathway Database (32) to identify metabolites and pathways associated with the 13 genes (*SI Appendix*, Fig. S16).

The 13 genes and 24 metabolites showing increased divergence between the transgenic and control mice converged within seven KEGG pathways (*SI Appendix*, Table S10). For three of these pathways, the HIF-1 signaling pathway, the pentose phosphate pathway, and carbon metabolism, metabolite concentration differences were significantly greater than expected by chance (permutations, $P < 0.05$) (Fig. 8*A*). The three pathways are closely related as the "carbon metabolism" pathway includes the pentose phosphate pathway and the TCA cycle, which is in turn linked to the HIF-1 signaling pathway (Fig. 8*B*).

We next analyzed metabolic differences between human and macaque brains. As the human samples were frozen after substantial postmortem delay (PMD), we first analyzed the effects of PMD by comparing three macaque samples that were intentionally collected with substantial PMD with macaque samples that were not (*SI Appendix*, Table S11). Of the 88 metabolites detected in the macaque and human brains, 21 were affected by PMD in at least one of the three samples analyzed (*SI Appendix*, Fig. S17 and Table S9) and were therefore excluded from further analyses.

For the remaining 67 metabolites, the differences between humans and macaques explained 26% of the total variation in metabolite concentrations. By comparison, the *GLUD2* genotype explains 2% of the metabolic variation in the mice (*SI Appendix*, Fig. S18). Despite the much greater overall metabolic differences between macaques and humans than between the transgenic and control mice, it is notable that in both cases, the greatest metabolic differences are seen in early development (*SI Appendix*, Figs. S15 and S19).

Of the 24 metabolites linked to the genes differentially expressed between *GLUD2* and control mice, 11 were among the

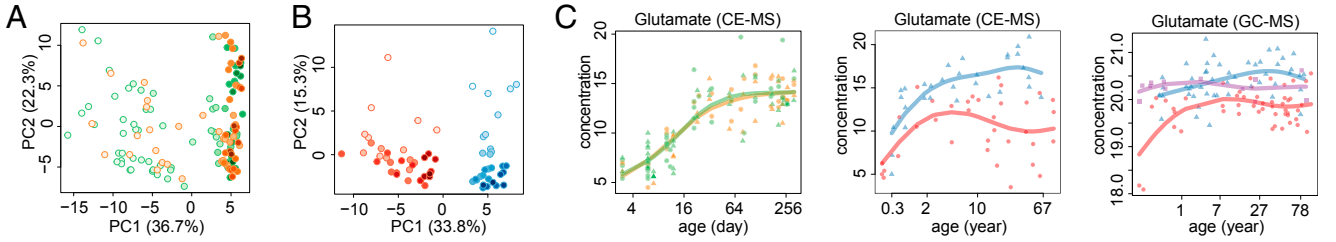

Fig. 6. Metabolite differences and glutamate concentration among genotypes/species. (*A* and *B*) The first two PCs of the principal component analysis, based on the concentrations of 110 metabolites detected in mouse frontal cortex and 88 metabolites detected in human and macaque prefrontal cortex. Each circle represents an individual. The colors indicate genotype/species information: orange, transgenic mice; green, control mice; red, humans; blue, rhesus macaques. Lighter shades of color indicate younger ages. (*C*) Glutamate concentration profiles measured in mice using CE-MS and in primates using CE-MS and GC-MS. The colors indicate genotype/species information as in *A* and *B*; purple, chimpanzees. Lines show spline curves fitted to concentration data points with four degrees of freedom.

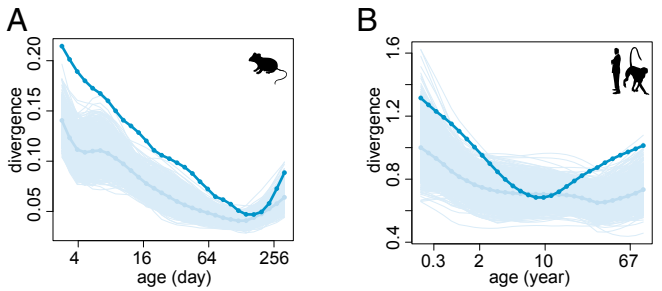

Fig. 7. Metabolome analyses of transgenic mice and primates. (*A*) The mean normalized metabolite concentration divergence between transgenic and control mice, based on the 24 metabolites sharing the same KEGG pathway as the 13 genes with expression affected by the *GLUD2* genotype (dark-blue curve) and the remaining 86 detected metabolites (light-blue curve). The colored area shows variation of the divergence estimates of the 86 metabolites obtained by bootstrapping 1,000 times. (*B*) The mean normalized metabolite concentration divergence between humans and macaques based on 11 of the 24 metabolites linked to the *GLUD2* genotype effect in mice (dark-blue curve) and the remaining 56 detected metabolites (light-blue curve). The colored area shows variation of the divergence estimates of the 56 metabolites obtained by bootstrapping 1,000 times.

67 metabolites that were detected and unaffected by PMD in humans and rhesus macaques (*SI Appendix*, Table S9). As in the mice, these 11 metabolites diverged more between humans and macaques than the other detected metabolites in early development (Fig. 7*B*). However, this effect was less pronounced in development and also apparent at later age, possibly due to fewer metabolites detected in the primates than in the mice, as well as differences in lifespan. Still, for 8 of the 11 metabolites detected in both primates and mice, the direction of concentration change coincided between humans and transgenic mice (permutation, $P = 0.08$) (*SI Appendix*, Fig. S20). Thus, *GLUD2* transgenic mice recapitulate some of the metabolic differences seen between human and macaque brains.

Discussion

GLUD2 originated as an evolutionary novel version of the glutamate dehydrogenase gene as a result of the retroposition of a *GLUD1* transcript in the common ancestors of humans and apes. We investigated the function of *GLUD2* by inserting the human *GLUD2* gene and surrounding sequences carrying putative regulatory elements into the mouse genome. We isolated two independent transgenic lines to exclude artifacts caused by the insertion of the human gene in the mouse genome. When analyzed in the frontal cortex, both lines displayed similar developmental *GLUD2* expression trajectories, closely resembling *GLUD2* expression in human prefrontal cortex during development.

In both transgenic lines, *GLUD2* affects the expression of 13 genes, 11 of which show similar ontogenetic expression profiles. These genes include several TFs, among which *DLX2*, *SOX4*, and *SOX11* play important roles in neuronal differentiation and neurogenesis (16, 18, 19, 25). Some of the previously described targets of these TFs show changes in expression in the transgenic mice that are correlated with expression changes of the corresponding TFs. Nine primate orthologs of the 13 mouse genes affected by *GLUD2*, including *DLX2*, *SOX4*, and *SOX11*, differ in expression between human and macaque ontogenesis in ways that mirror the changes seen in the mice. Similarly, primate orthologs of target genes of *DLX2*, *SOX4*, and *SOX11* show expression differences correlating with those of these TFs. These results illustrate that the introduction of *GLUD2* into the mouse genome induces effects paralleling evolutionary differences between primate species that carry a *GLUD2* gene and those that do not. This adds to a mounting amount of evidence suggesting that human-specific variants of genes such as *FOXP2*, *ASPM*, *EDAR*, *SRGAP2*, and *CMAH* can be fruitfully studied in mouse models (33).

The gene expression changes induced by *GLUD2* in the mice and the differences between humans and macaques were restricted to early development and were not observed past the first 2 wk postpartum in the mice or the first 2 y of life in humans.

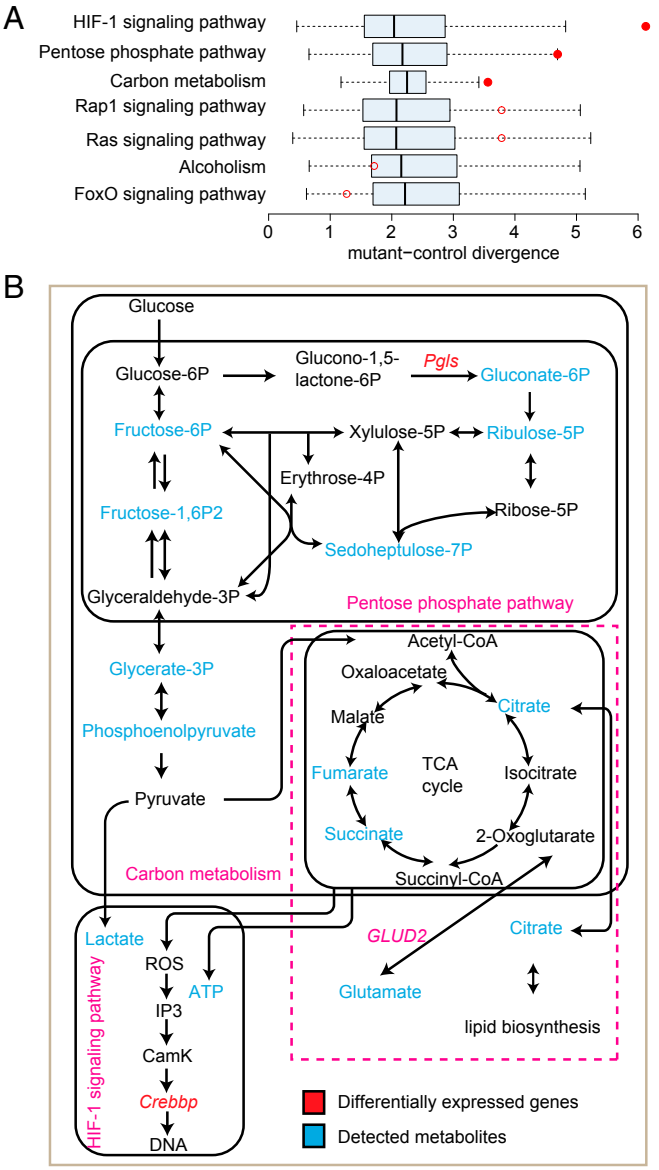

Fig. 8. Pathway analysis of the *GLUD2* genotype effect on mouse metabolome. (*A*) Metabolite concentration divergence between transgenic and control mice, based on metabolites in a pathway (red circles). The boxplots show metabolite concentration divergence between transgenic and control mice, calculated by sampling 1,000 times the same number of metabolites as detected in a given pathway from the bulk of remaining detected metabolites. Each pathway contains at least two detected metabolites and at least one gene differentially expressed between transgenic and control mice. The filled circles show the pathways with significantly greater metabolic divergence than expected by chance. (*B*) The schematic representation of the three KEGG pathways showing significantly greater metabolite concentration divergence between transgenic and control mice: HIF-1 signaling pathway, pentose phosphate pathway, and carbon metabolism. Detected metabolites are shown in blue and genes with expression affected by the *GLUD2* genotype are shown in red. Dashed pink rectangle delineates the pathway affected by *GLUD2* overexpression in *IDH1*-mutant glioma cells.

This suggests that *GLUD2* mainly functions during brain growth and early development, a notion that is in apparent contradiction to the idea that *GLUD2* has a major role in the metabolism of the neurotransmitter glutamate.

The metabolic data lend support to the notion that *GLUD2* mainly influences aspects of metabolism different from glutamate recycling. We detect no effects on glutamate concentrations in the mouse frontal cortex and our previous results show that glutamate concentrations do not differ between macaques and chimpanzees throughout prefrontal cortex developmental and adult stages, even though chimpanzees carry a *GLUD2* gene. Instead, metabolic differences between $GLUD2^{+}$ and control mice center on the metabolic pathways surrounding the TCA cycle. In agreement with this, it has been reported that overexpression of human *GLUD2*, but not *GLUD1*, in mutant murine glioma progenitor cells results in shunting of carbon into lipid biosynthesis via the TCA cycle (12). Our data further show that metabolic differences between $GLUD2^{+}$ and control mice, as well as between humans and macaques, are particularly prevalent at early developmental stages characterized by rapid brain growth. Given that lipids comprise more than 50% of the brain's dry weight (34), we speculate that *GLUD2* may support the rapid growth of the large ape and human brains by enhancing lipid biosynthesis.

Materials and Methods

A detailed description of materials and methods is provided in *SI Appendix*. Briefly, a human bacterial artificial chromosome containing the *GLUD2* gene (RP11-610G22) was linearized by NotI and injected into the male pronucleus of C57BL/6 mice to construct $GLUD2^{+/-}$ transgenic mice. Transgenic animals were identified by PCRs targeting the 5′-end, the coding region, and the 3′-end of the gene. For metabolome analysis, CE-MS measurements were conducted in frontal cortex samples of 81 control and 56 transgenic mice and prefrontal cortex samples of 35 humans and 34 rhesus macaques. Transcriptome analysis was conducted in the frontal cortex of 29 control and 30 transgenic mice used for metabolite profiling on the Illumina platform, as well as in prefrontal cortex of 38 humans and 40 rhesus macaques (14). Written consent for the use of human tissues for research was obtained from all donors or their next of kin. Use of human autopsy tissue is considered nonhuman subject research and is institutional review board exempt under NIH guidelines. Biomedical Research Ethics Committee of Shanghai Institutes for Biological Sciences completed the review of the use and care of the animals in the research project (approval ID: ER-SIBS-260802P).

ACKNOWLEDGMENTS. We thank Rowina Voigtländer and the staff of the Max Planck Institute for Evolutionary Anthropology's animal facility for expert mouse care; Wulf Hevers and Ines Bliesener for help in preparation of mouse samples; Dr. H. R. Zielke and Dr. J. Dai for providing the human samples; C. Lian, H. Cai, and X. Zheng for providing the macaque samples; J. Boyd-Kirkup for editing the manuscript; Z. He for assistance; and all members of the Comparative Biology Group in Shanghai for helpful discussions and suggestions. This study was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development Brain and Tissue Bank for Developmental Disorders; the Netherlands Brain Bank; the Chinese Brain Bank Center; the Suzhou Drug Safety Evaluation and Research Center; the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant XDB13010200 to P.K.); the National Natural Science Foundation of China (Grants 91331203 and 31420103920 to P.K.); the National One Thousand Foreign Experts Plan (Grant WQ20123100078 to P.K.); the National Natural Science Foundation of China (Grant 31401065 to S.G.); the Max Planck Society; and the Paul G. Allen Family Foundation (S.P.).

1. Shashidharan P, et al. (1994) Novel human glutamate dehydrogenase expressed in neural and testicular tissues and encoded by an X-linked intronless gene. *J Biol Chem* 269(24):16971–16976.
2. Fonnum F (1984) Glutamate: A neurotransmitter in mammalian brain. *J Neurochem* 42(1):1–11.
3. Orrego F, Villanueva S (1993) The chemical nature of the main central excitatory transmitter: A critical appraisal based upon release studies and synaptic vesicle localization. *Neuroscience* 56(3):539–555.
4. Burki F, Kaessmann H (2004) Birth and adaptive evolution of a hominoid gene that supports high neurotransmitter flux. *Nat Genet* 36(10):1061–1063.
5. Dupanloup I, Kaessmann H (2006) Evolutionary simulations to detect functional lineage-specific genes. *Bioinformatics* 22(15):1815–1822.
6. Plaitakis A, Spanaki C, Mastorodemos V, Zaganas I (2003) Study of structure-function relationships in human glutamate dehydrogenases reveals novel molecular mechanisms for the regulation of the nerve tissue-specific (GLUD2) isoenzyme. *Neurochem Int* 43(4-5):401–410.
7. Kanavouras K, Mastorodemos V, Borompokas N, Spanaki C, Plaitakis A (2007) Properties and molecular evolution of human GLUD2 (neural and testicular tissue-specific) glutamate dehydrogenase. *J Neurosci Res* 85(5):1101–1109.
8. Rosso L, Marques AC, Reichert AS, Kaessmann H (2008) Mitochondrial targeting adaptation of the hominoid-specific glutamate dehydrogenase driven by positive Darwinian selection. *PLoS Genet* 4(8):e1000150.
9. Kotzamani D, Plaitakis A (2012) Alpha helical structures in the leader sequence of human GLUD2 glutamate dehydrogenase responsible for mitochondrial import. *Neurochem Int* 61(4):463–469.
10. Shashidharan P, Plaitakis A (2014) The discovery of human of GLUD2 glutamate dehydrogenase and its implications for cell function in health and disease. *Neurochem Res* 39(3):460–470.
11. Plaitakis A, Latsoudis H, Spanaki C (2011) The human GLUD2 glutamate dehydrogenase and its regulation in health and disease. *Neurochem Int* 59(4):495–509.
12. Chen R, et al. (2014) Hominoid-specific enzyme GLUD2 promotes growth of IDH1R132H glioma. *Proc Natl Acad Sci USA* 111(39):14217–14222.
13. Bryk J (2009) How to Make an Ape Brain: A Transgenic Mouse Model of Brain Glutamate Metabolism in Humans and Apes. PhD thesis (Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany).
14. He Z, Bammann H, Han D, Xie G, Khaitovich P (2014) Conserved expression of lincRNA during human and macaque prefrontal cortex development and maturation. *RNA* 20(7):1103–1111.
15. Tabas-Madrid D, Nogales-Cadenas R, Pascual-Montano A (2012) GeneCodis3: A non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res* 40(Web Server ssue):W478–W483.
16. Shim S, Kwan KY, Li M, Lefebvre V, Sestan N (2012) Cis-regulatory control of corticospinal system development and evolution. *Nature* 486(7401):74–79.
17. Bhattaram P, et al. (2010) Organogenesis relies on SoxC transcription factors for the survival of neural and mesenchymal progenitors. *Nat Commun* 1:9.
18. Mu L, et al. (2012) SoxC transcription factors are required for neuronal differentiation in adult hippocampal neurogenesis. *J Neurosci* 32(9):3067–3080.
19. Wang Y, Lin L, Lai H, Parada LF, Lei L (2013) Transcription factor Sox11 is essential for both embryonic and adult neurogenesis. *Dev Dyn* 242(6):638–653.
20. Jankowski MP, Cornuet PK, McIlwrath S, Koerber HR, Albers KM (2006) SRY-box containing gene 11 (Sox11) transcription factor is required for neuron survival and neurite growth. *Neuroscience* 143(2):501–514.
21. Hide T, et al. (2009) Sox11 prevents tumorigenesis of glioma-initiating cells by inducing neuronal differentiation. *Cancer Res* 69(20):7953–7959.
22. Le TN, et al. (2007) Dlx homeobox genes promote cortical interneuron migration from the basal forebrain by direct repression of the semaphorin receptor neuropilin-2. *J Biol Chem* 282(26):19071–19081.
23. Tanaka S, et al. (2004) Interplay of SOX and POU factors in regulation of the Nestin gene in neural primordial cells. *Mol Cell Biol* 24(20):8834–8846.
24. Kim HD, et al. (2011) Class-C SOX transcription factors control GnRH gene expression via the intronic transcriptional enhancer. *Mol Endocrinol* 25(7):1184–1196.
25. Paina S, et al. (2011) Wnt5a is a transcriptional target of Dlx homeogenes and promotes differentiation of interneuron progenitors in vitro and in vivo. *J Neurosci* 31(7):2675–2687.
26. Iyer AK, Miller NL, Yip K, Tran BH, Mellon PL (2010) Enhancers of GnRH transcription embedded in an upstream gene use homeodomain proteins to specify hypothalamic expression. *Mol Endocrinol* 24(10):1949–1964.
27. Colasante G, et al. (2008) Arx is a direct target of Dlx2 and thereby contributes to the tangential migration of GABAergic interneurons. *J Neurosci* 28(42):10674–10686.
28. de Melo J, et al. (2008) Dlx2 homeobox gene transcriptional regulation of Trkb neurotrophin receptor expression during mouse retinal development. *Nucleic Acids Res* 36(3):872–884.
29. Bergsland M, Werme M, Malewicz M, Perlmann T, Muhr J (2006) The establishment of neuronal properties is controlled by Sox4 and Sox11. *Genes Dev* 20(24):3475–3486.
30. Miller JA, et al. (2014) Transcriptional landscape of the prenatal human brain. *Nature* 508(7495):199–206.
31. Fu X, et al. (2011) Rapid metabolic evolution in human prefrontal cortex. *Proc Natl Acad Sci USA* 108(15):6181–6186.
32. Jewison T, et al. (2014) SMPDB 2.0: Big improvements to the Small Molecule Pathway Database. *Nucleic Acids Res* 42(Database issue):D478–D484.
33. Enard W (2014) Mouse models of human evolution. *Curr Opin Genet Dev* 29:75–80.
34. Piomelli D, Astarita G, Rapaka R (2007) A neuroscientist's guide to lipidomics. *Nat Rev Neurosci* 8(10):743–754.

Supporting Information

SI Materials and Methods

Construction of GLUD2$^{+/-}$ transgenic mice

A human bacterial artificial chromosome containing the *GLUD2* gene (RP11-610G22) was linearized by *NotI* and injected into the male pronucleus of C57Bl/6 mice. Transgenic animals were identified by PCRs targeting the 5'-end, the coding region, and the 3'-end of the gene. Both lines show inheritance patterns consistent with single autosomal locations of the *GLUD2* transgenes. The two mouse lines were maintained under pathogen-free conditions as crosses between hemizygous lines by crosses with non-transgenic mice. All comparisons were between transgenic and non-transgenic littermates.

Samples

Human samples were obtained from the NICHD Brain and Tissue Bank for Developmental Disorders at the University of Maryland, USA, the Netherlands Brain Bank, Amsterdam, Netherlands and the Chinese Brain Bank Center (CBBC), Wuhan, China. Written consent for the use of human tissues for research was obtained from all donors or their next of kin. All subjects were defined as healthy controls by forensic pathologists at the corresponding brain bank. All subjects suffered sudden death with no prolonged agonal state. According to the protocol of the CBBC, use of human autopsy tissue is considered non human-subject research and is IRB exempt under NIH guidelines. Rhesus macaque samples were obtained from the Suzhou Experimental Animal Center, China. All rhesus macaques used in this study suffered sudden deaths for reasons other than their participation in this study and without any relation to the tissue used. Prefrontal cortex (PFC) dissections were made from the frontal part of the superior frontal gyrus. For all samples we took special care to dissect gray matter only. In order to identify the metabolites affected by PMD in primates, we collected three additional tissue samples from two rhesus macaques dissected 4-6 hours after death.

CE-MS measurements

CE-MS measurements were conducted in frontal cortex samples of 81 control and 56 transgenic mice and PFC samples of 35 humans and 34 rhesus macaques. For each sample, metabolites were extracted from the frozen tissue powder by 1ml methanol containing 20µM each of L-Methionine sulfone, 2-Morpholinoethanesulfonic acid, monohydrate and sodium d-camphor-10-sulfonic acid. Then, 500µl of lysate was transferred to an Eppendorf tube containing 500µl chloroform and 200µl of Milli-Q water. 300µl of the aqueous phase, after 30 seconds of vortexing and 15 min of centrifugation at 4°C, was transferred to an ultrafiltration tube (Milipore). Concentrating the filtered liquid to complete dryness was then performed in a speed vacuum for 3 hours at 35°C. Just before the CE-MS analyses, the dried samples were mixed with 100µl of Milli-Q water containing 100µM each of 3-aminopyrrolidine and trimesate, and filtered with an ultrafiltration tube (Milipore) at 9,100×g for 2h at 4°C. 7µl of filtrate was used for the CE-MS analyses. CE-TOF-MS (Agilent Technologies) was then used to detect both cationic metabolites and anionic metabolites. The

instrumentation and measurement conditions used for CE-TOF-MS were according to (1).

The in-house software MasterHands was used to perform peak detection, time alignment, and peak area integration. Concentrations of each metabolite in the samples were calculated based on the comparison of peak area normalized by internal standards' in the sample and external standard mixture. Metabolites with concentrations detected in more than half of the samples were included for following analysis.

RNA-seq measurements

For the mouse model, tissue samples from the frontal cortex of 29 control and 30 transgenic mice used for metabolite profiling, were selected for RNA-sequencing. The poly(A)+ RNA fraction was isolated and processed for sequencing using standard cDNA library preparation protocol. Library sequencing was carried out using the Illumina HiSeq 2000 platform with a sequence length of 100 bp. In the order of 10 – 14 million reads were generated for each sample.

In order to map the resulting sequencing reads, we complemented the mouse reference genome with the human genomic region inserted into transgenic mice to construct a custom reference genome. We then used the "STAR" (2) algorithm to map the reads, allowing at most one mismatch and requiring the number of matched bases out of 100 bases to be equal or greater than 95. Reads mapped to multiple genomic locations or containing non-canonical junctions were excluded from downstream analysis. Based on the mapping alignments and mouse GENCODE vM1 gene annotation, we calculated gene expression level as reads per kilobase per million mapped reads (RPKM). Only genes with RPKM > 0 in at least half of the samples and with sum of RPKM in all the samples greater than 5 were used in downstream analysis.

For primates, time series gene expression data collected in PFC of 38 humans (from 2 days to 61 years) and 40 rhesus macaques (from 13.6 weeks post-conception to 21 years) measured using Illumina RNA-seq platform was taken from (3). Based on the read alignments and GENCODE v19 annotation, we calculated the expression level of each gene as RPKM. We only retained genes with RPKM > 0 in at least half of the samples and with the sum of RPKM in all the samples greater than 5. Another human time series gene expression data, collected in PFC samples across 13 developmental stages, from embryo to adulthood, was downloaded from BrainSpan, an atlas of the developing human brain (http://www.developinghumanbrain.org).

PCA

For the PCA analysis, the "prcomp" function in the R package "stats" was used.

Age scale

To produce a more uniform age distribution, all analyses conducted in the mouse model were based on the $\log_2$-transformed sample ages. For analyses performed in primates, we took gestation time into account by adding 280 days and 165 days to the ages of humans and rhesus macaques, respectively. We then normalized lifespan differences between humans and macaques by multiplying macaque ages by three, based on the maximum lifespan difference between species: 105 years for humans and

35 years for macaques. The resulting ages were then $\log_2$-transformed to produce a more uniform age distribution.

Age test and differential expression test

To test the effect of age on the gene expression level, we used polynomial regression models with age as predictor and expression level as response, described in (4). Briefly, for each gene, we chose the best polynomial regression model, based on families of polynomial regression models and "adjusted r^2" criterion. The significance of the chosen model was then estimated using F-test. Transgenic mice and control mice were tested independently. Age-dependent genes were defined as those with age-test *p*-values less than 0.05 after Benjamini Hochberg correction in at least one of the genotypes. The false discovery rate (FDR) was estimated by 1,000 random permutations of ages across samples.

To detect whether transgenic mice and control mice have different profiles of expression change with age, we used ANCOVA, as described in (4). Briefly, for each gene, based on the polynomial regression model obtained by the above-described age-test, we tested if such a regression model with genotype-specific parameters is significantly better than the model with common parameters for both genotypes, given the expression-age distribution of the two genotypes. The null model (with no genotype-specific parameters) and alternative models were compared by F-test. The differentially expressed genes between transgenic mice and control mice were defined as those with *p*-values less than 0.05 after Benjamini Hochberg correction.

The permutations for the differential expression test was conducted by dividing the age-range into 10 sections and randomly permuting the genotype assignments across samples within each section to preserve the age-structure in the data, 1,000 times. The permutation *p*-value was calculated as the frequency of the cases when the number of differentially expressed genes between transgenic and control mice in a random permutation was equal or greater than the actual one. The FDR was estimated as the ratio of the median of the permutation distribution to the actual one.

Gene expression/metabolite concentration variance explained by factors

The gene expression/metabolite concentration variance explained by genotype/species, age, line (for mice) and sex were estimated using the method described in (4). In brief, for each gene/metabolite, separate regression models were fitted for age, genotype/species, line or sex, based on the mouse/primate dataset. For genotype/species, line, or sex, linear regression models were applied, whereas for age, third degree polynomial models based on age-test were used. The proportion of total variance explained by each factor was defined as one minus the ratio of residual sum of squares from the regression model and residual sum of squares from the null model (variance multiplied by n-1, where n is the sample size). The significance of the mean proportion of variance explained across all detected genes/metabolites was then estimated by randomly permuting the factor (genotype/species, age, line, or sex) assignments across samples, 1,000 times. The permutation *p*-values were calculated as the frequency of permutations where the permutated mean proportion was equal or greater than the real mean proportion.

Expression/concentration curve fitting

For genes/metabolites, a spline curve was fitted to the expression/concentration data points for each genotype or species. Cubic spline regression with four degrees of freedom was used for spline curve fitting.

GLUD2 expression profile in humans and transgenic mice

As the sequences of *GLUD1* and *GLUD2* genes in the human genome are very similar, we remapped raw reads from humans from (3) to the human hg19 reference genome using the program "TopHat2" (5) with stringent criteria, allowing no mismatches and mapping every read in all mapping steps. Only uniquely mapped reads were considered for the *GLUD2* expression level estimation. RPKM was calculated for *GLUD2* based on the read alignments and GENCODE v19 annotation. To estimate the similarity of *GLUD2* expression profiles between humans and transgenic mice, Pearson correlation coefficient was calculated based on two vectors of 30 points interpolated from spline curves fitted to human *GLUD2* and transgenic mouse *GLUD2* expression levels with four degrees of freedom.

Bidirectional transcription of GLUD2 and the upstream lncRNA

To test whether *GLUD2* and the upstream lncRNA share the same transcription initiation locus, deepCAGE data contained within the R package "FANTOM3and4CAGE" were used to count CAGE tags at both strands of the *GLUD2* promoter region.

Divergence calculation

In order to determine the stages when the differences between genotypes/species occurred, we calculated the divergence between genotypes/species along the lifespan. First, for each gene/metabolite and for each genotype/species, we interpolated expression/concentration values at 30 equally distributed points along the age range. Interpolation was performed using cubic spline regression with four degrees of freedom. We then defined the divergence between genotypes/species for each gene/metabolite at each of the 30 age points as follows:

$$D_{ij} = \frac{|C_{ij} - M_{ij}|}{E_i}$$

where D_{ij} is the divergence for gene/metabolite i at age point j, C_{ij} and M_{ij} are the interpolated expression/concentration values of each genotypes/species for gene/metabolite i at age point j,E_i is the mean expression/concentration value of these two genotypes/species across all age points for gene/metabolite i. E_i is used to remove the effect of absolute expression/concentration value of gene/metabolite i on the divergence estimation.

For the divergence of a gene/metabolite set, the divergence at each age point was calculated as the mean divergence of all the genes/metabolites within the set at this age point. The variance of the divergence was estimated by either resampling among the gene/metabolite set with replacement 1,000 times (bootstrapping) or randomly sampling genes/metabolites with the same number as that of the other set 1,000 times (sub-sampling).

The significance of the divergence between humans and macaques for the 9 genes

with expression affected by *GLUD2* was conducted by random sub-sampling of 9 genes from the remaining genes expressed in human and macaque PFC 1,000 times. *P*-value was defined as the frequency of cases when the divergence obtained in the sub-sampling showed the same amplitude and pattern as observed for the 9 genes identified in the *GLUD2* mouse model (Pearson correlation $r > 0.9$, max divergence in sub-sampling $\geq$ actual max divergence).

Hierarchical clustering analysis

To identify co-expressed modules based on the gene expression levels in the mouse model, we used unsupervised hierarchical clustering. Since most genes had no significant differences between the two genotypes, the hierarchical clustering was based on the expression levels from control mouse samples. Pearson correlation coefficient (r) between the expression levels of pairs of genes was used as the distance measurement (i.e. 1-r). We chose the method "complete" for the clustering and cut the tree at a height determined by manual examination of cluster profiles (h=1.5). Based on all the 15,663 age-dependent genes, we identified 28 co-expressed modules.

Out of 13 genes differentially expressed between transgenic mice and control mice, 11 fell into module 1. The significance of this observation was estimated by randomly replacing the differentially expressed genes with other age-dependent genes 1,000 times. *P*-value was determined as the frequency of events in which the number of genes classified into module 1in a random replacement was greater or equal to 11.

Developmental time-shift analysis

To estimate the shift in timing of expression between genotypes/species, we used DTW-Significance, a modified dynamic time warping algorithm (6). Briefly, this algorithm finds the best alignment between two time series expression data, making the differences between them minimal. The algorithm then reports the shift in timing of expression – the differences between the corresponding ages at which the expressions of the two data correspond to each other. Here, we used this algorithm to find the matched age points at which expression in transgenic mice/humans correspond to that in control mice/rhesus macaques. The alignment was done between the expression levels at 20 equally distributed time points in transgenic mice/rhesus macaques and expression levels at 20 time points sampled from 40 equally distributed time points in control mice/humans. The time points covered the whole lifespan of each species/genotypes, from the birth point to the maximum point among our samples. In order to make the time-shift estimation more accurate, we focused on genes satisfying the following criteria: (1) showing significant expression change with age; (2) having positive correlation between genotypes/species (Pearson correlation $P < 0.05$ and $r > 0$).

For the time-shift analysis in mice, the average time-shift at each time point across the 11 module 1 genes with expression affected by *GLUD2* genotype, and across the remaining genes satisfying the above criteria, was calculated. To estimate the variance of the time-shift of the remaining genes, we randomly sub-sampled 11 genes from the remaining genes, 1,000 times.

For the time-shift analysis in primates, the average time-shift at each time point across the 7 genes showing the similar expression and the similar divergence pattern in primates and mice was estimated. The Pearson correlation coefficient was then

calculated based on the time-shift of each pair of the 7 genes, which resulted in 21 correlated gene pairs ($r > 0.9$). To estimate the significance of the observation, we randomly sub-sampled 7 genes from the remaining expressed genes satisfying the above criteria in primates 1,000 times. *P*-value was defined as the frequency of cases when the number of correlated gene pairs in a random sub-sampling was equal or greater than 21.

Functional enrichment analysis

We identified overrepresented Gene Ontology (GO) terms using "GeneCodis". The gene list was the 11 module 1 genes and the reference list was all age-dependent genes. We used default statistical parameters – hypergeometric statistical test and FDR *p*-value correction. GO terms with corrected *p*-value less than 0.05 were considered to be enriched.

Developmental expression pattern in transgenic mice and humans

In order to estimate the similarity of the developmental expression profiles between transgenic mice and humans for the 9 genes with expression affected by *GLUD2* genotype, we calculated the Pearson correlation coefficient between two vectors of 30 points which were interpolated from spline curves fitted to transgenic mice and human expression data points with four degrees of freedom. 7 among the 9 genes showed consistent expression profiles between transgenic mice and humans ($r > 0.9$). To estimate the significance of the observation, we randomly sub-sampled 9 genes from the remaining expressed genes in primates 1,000 times. *P*-value was defined as the frequency of the cases when the number of genes showing consistent expression profiles between transgenic mice and humans ($r > 0.9$) in a sub-sampling was equal or greater than 7.

SOX4, SOX11, DLX2 target gene analysis

We manually collected targets of these three TFs from "PubMed" (http://www.ncbi.nlm.nih.gov/pubmed/). Fifteen known target genes were found to be expressed in our mouse data and 14 of them were expressed in our primate data.

To test the downstream effect of these TFs, we calculated the Pearson correlation coefficient based on the divergence of the transcription factors and that of their targets. Transcription factor–target pairs with correlation coefficient greater than 0.5 ($r > 0.5$) were classified as correlated ones. In order to estimate the significance of the correlation, we randomly sub-sampled gene sets with the same number as that of targets and then calculated the correlation based on the divergence of the transcription factors and that of the random targets, 1,000 times. *P*-value was defined as the frequency of the events when the number of correlated transcription factor-random target pairs was equal or greater than the number of correlated transcription factor-real target pairs.

Glutamate measurements using GC-MS

Metabolite measurements conducted in PFC of 50 humans, 12 chimpanzees and 49 rhesus macaques using GC-MS were downloaded from (7). Concentrations of each compound were normalized by internal standard and log2-transformed. Before log2-transformation, the metabolite/internal standard ratios were multiplied by 10,000 to avoid negative values. Compounds with concentrations detected in less than 35%

of the samples were removed. The remaining missing values were imputed using the R package "impute". Ages of the three species were normalized to the same maximal lifespan and then were scaled by fourth root to produce uniform distributions. Spline curves with four degrees of freedom were fitted to the concentration of glutamate for the three species, respectively.

Metabolite analyses in the mice

For the genes with expression affected by the *GLUD2* genotype in the mouse model, we first identified the pathways where those genes were located based on KEGG pathway annotation (http://www.kegg.com). We then collected metabolites that were detected in our dataset from these identified pathways, which resulted in 24 such metabolites. The divergence of the 24 metabolites and the remaining detected metabolites was estimated based on the method introduced above.

In addition to KEGG pathway annotation, we used Small Molecule Pathway Database (SMPDB, http://smpdb.ca/). The same procedures as above were performed to obtain the divergence of the metabolites sharing the same SMPDB pathway as the genes with expression affected by *GLUD2* genotype and the divergence of the remaining detected metabolites.

KEGG pathway analysis

In the mouse model, the 13 genes differentially expressed between transgenic mice and control mice and the 110 detected metabolites were mapped to KEGG pathways. Pathways that contained at least one differentially expressed gene and at least two detected metabolites were considered for following analysis. Based on this criterion, we obtained 7 KEGG pathways. For each of these pathways, for each metabolite within it, the sum divergence across all age points was calculated. The average of the sum divergence across all the metabolites within this pathway was used to represent the divergence of this pathway. To estimate the significance of each pathway, we replaced the detected metabolites within this pathway with the same number of metabolites randomly sampled from the remaining detected metabolites 1,000 times. *P*-value was determined by the frequency of events when the divergence in a random sub-sampling was equal or greater than the one obtained in the actual pathway.

Identification of metabolites affected by PMD

In order to test the effect of PMD in metabolite concentration analysis in primates, we compared the concentration levels from two macaques individuals (one of them has two replicates) collected with substantial PMD with those from the remaining normal individuals. For each metabolite, we used a spline curve to fit the concentration levels from normal macaque samples with four degrees of freedom. The differences between the real concentration levels and the interpolated values from the spline curve were defined as residuals. If the difference between the real concentration level from one of the three macaque samples with substantial PMD, and the value interpolated from the above spline curve, fell outside of the 1.96*standard deviation of the residuals, this metabolite was classified as PMD affected metabolites.

Metabolite analyses in primates

Among the 88 metabolites detected in primates, 67 were not affected by PMD. 11 of the 67 metabolites can be linked to the 24 metabolites that shared the same KEGG

pathways as the differentially expressed genes in the mouse model. The divergence between humans and rhesus macaques of the 11 metabolites and the remaining detected metabolites was calculated based on the method described above. For comparison of the metabolite abundance change between primates and mice, we calculated average metabolite concentration difference between human and macaque or between control and transgenic mice time series for each of the 11 shared compounds normalized by the average concentration of the compound. We tested the significance of the concentration change direction agreement between primates and mice by sub-sampling 11 compounds from the remaining mouse and primate data.

SI Figures

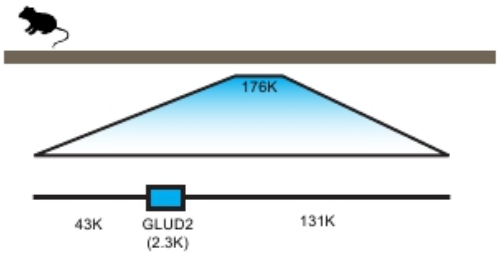

Fig. S1. Schematic representation of the *GLUD2* transgenic region in the mouse genome. A 176 kb long human genomic region carrying the mono-exonic *GLUD2* gene, as well as 43 kb of upstream sequences and 131 kb of downstream sequence, were inserted into the mouse genome.

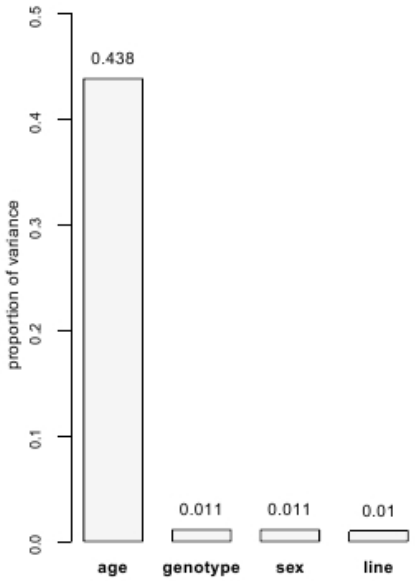

Fig. S2. The proportion of gene expression variation explained by age, genotype, sex and line factors in the mouse brain dataset. A regression model with age, genotype, sex or line factors was fitted to the expression level of each of the 18,610 genes detected in mouse frontal cortex. The proportion of variance explained by each factor was calculated as one minus the ratio of residual sum of squares from the regression model and residual sum of squares from the null model. The mean proportion of variance explained by each factor across the 18,610 genes is shown above the bars.

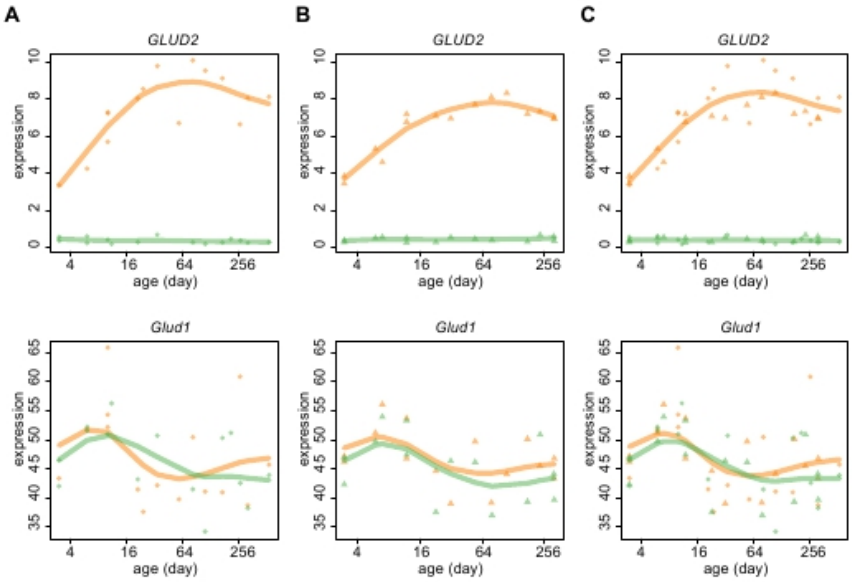

Fig. S3. *GLUD2* and *Glud1* expression profiles in the mouse model. The expression profiles of *GLUD2* and *Glud1* in the transgenic mice of "line a" and their control littermates (A), in the transgenic mice of "line b" and their control littermates (B), and in the transgenic mice of both lines and their control littermates (C). Each point represents an individual. The colors indicate genotype information: orange – transgenic mice; green – control mice. The symbols indicate mice lines: circles – line a; triangles – line b. The lines show spline curves fitted to expression data points with four degrees of freedom. The expression levels (y-axis) are measured as RPKM (reads per kilobase per million mapped reads).

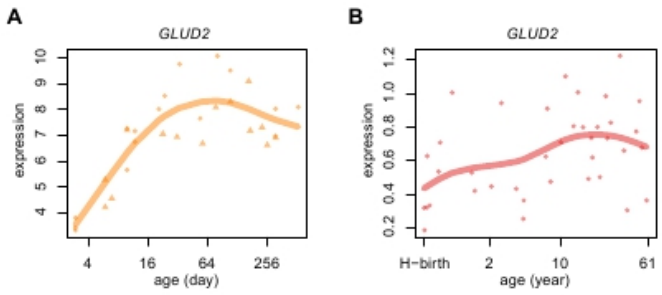

Fig. S4. *GLUD2* expression profile in transgenic mice and humans during postnatal development. (A) *GLUD2* expression in the frontal cortex of transgenic mice. (B) *GLUD2* expression in the human PFC. Each point represents an individual (orange – transgenic mice; red – humans). The symbols indicate mice lines: circles – line a; triangles – line b. The lines are spline curves fitted to *GLUD2* expression data points with four degrees of freedom.

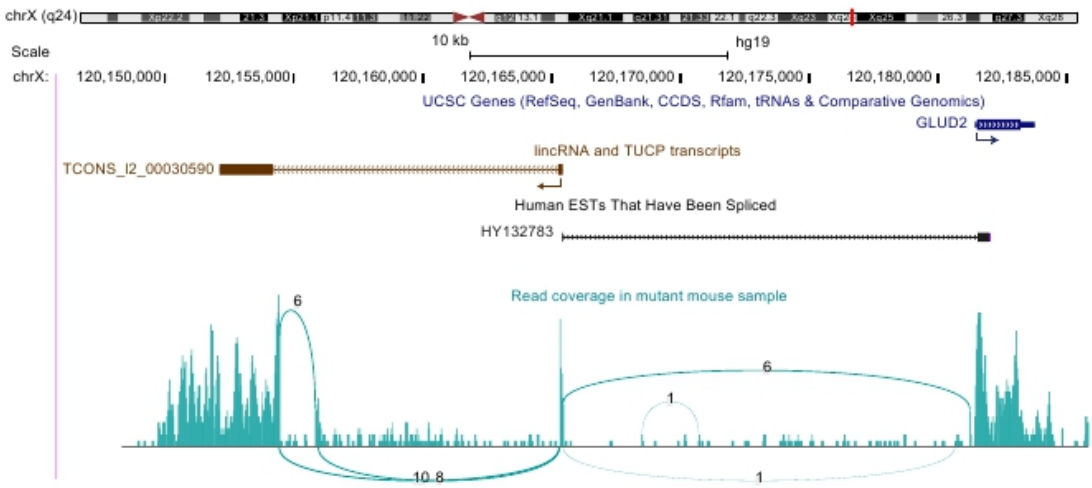

Fig. S5. A schematic representation of the *GLUD2* gene and the co-expressed lncRNA, located within the inserted human genomic region. The dark blue box shows the exon of *GLUD2* gene, and brown boxes show exons of the co-expressed lncRNA from Human Body Map LincRNAs. This lncRNA is located on the opposite strand of the human genomic region upstream of the *GLUD2* transcription start site. The black box represents the human ESTs, and blue peaks represent the read coverage based on one transgenic mouse sample. The bottom track is based on RNA-seq data from the transgenic mouse brain. Blue arcs and the numbers above them show the junction reads that connect exons. The human ESTs, and the junction reads in the transgenic mouse sample, indicate a potential exon of this lncRNA near the transcription start site of the *GLUD2* gene, suggesting a bidirectional promoter regulating the transcription of the *GLUD2* gene and this lncRNA. The annotations of *GLUD2*, this lncRNA and human ESTs are from UCSC Genome Browser (http://genome.ucsc.edu/).

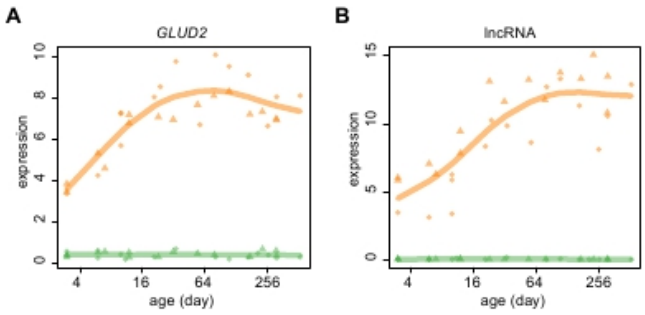

Fig. S6. The expression profiles of *GLUD2* gene and the upstream antisense lncRNA in mouse brain. (A and B) *GLUD2* and lncRNA expression in the frontal cortex of transgenic mice during postnatal development. Each point represents an individual (orange – transgenic mice; green – control mice). The symbols indicate mice lines: circles – line a; triangles – line b. The lines are spline curves fitted to *GLUD2* gene and the lncRNA expression data points with four degrees of freedom.

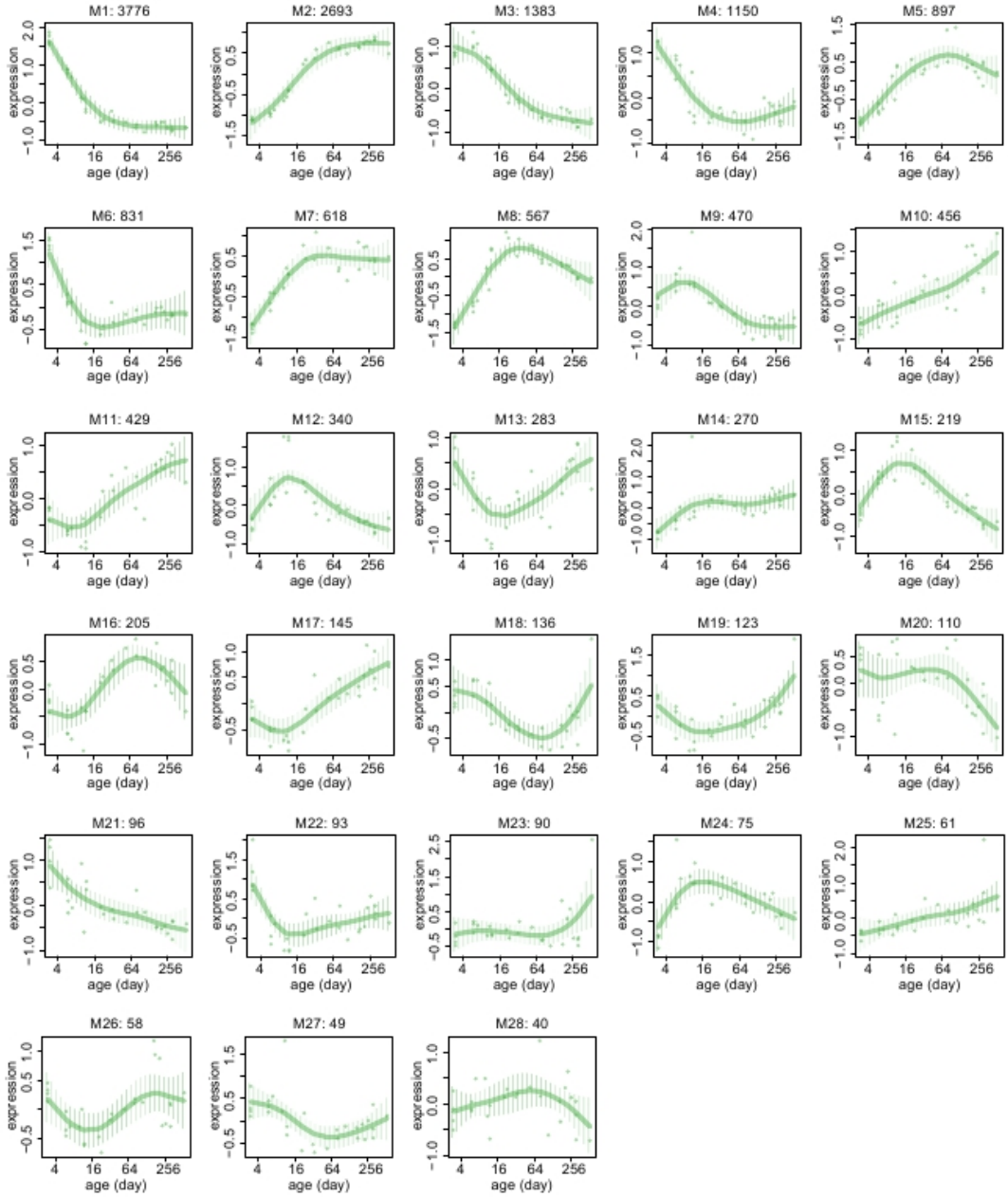

Fig. S7. Co-expressed modules obtained by unsupervised hierarchical clustering of expression profiles of the 15,663 age-dependent genes in mouse brain. The expression levels of each of the 15,663 genes were standardized to mean = 0 and standard deviation (SD) = 1. Points show the average expression level across the genes within a module. The lines show spline curves fitted to the average expression data points with four degrees of freedom. Error bars show the SD across the genes in a module. The numbers above each of the panels indicate the co-expressed module size. 11 of the 13 genes with expression affected by *GLUD2* genotype fell into module 1.

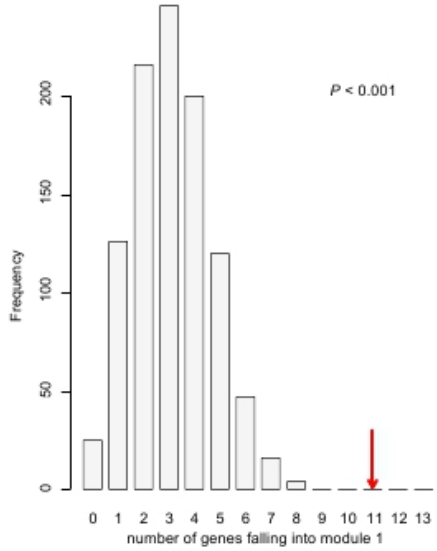

Fig. S8. The occurrence of age-dependent genes with expression affected by *GLUD2* genotype in module 1. The red arrow shows the number of genes falling into module 1 among the 13 genes with expression affected by the *GLUD2* genotype. Gray bars show the distribution of age-dependent gene numbers assigned into module 1, obtained by sub-sampling 13 genes from the remaining 15,650 age-dependent genes whose expressions were not affected by *GLUD2* genotype, 1,000 times. *P*-value is defined as the frequency of cases when the number of genes falling into module 1 in a sub-sampling was equal or greater than the actual one.

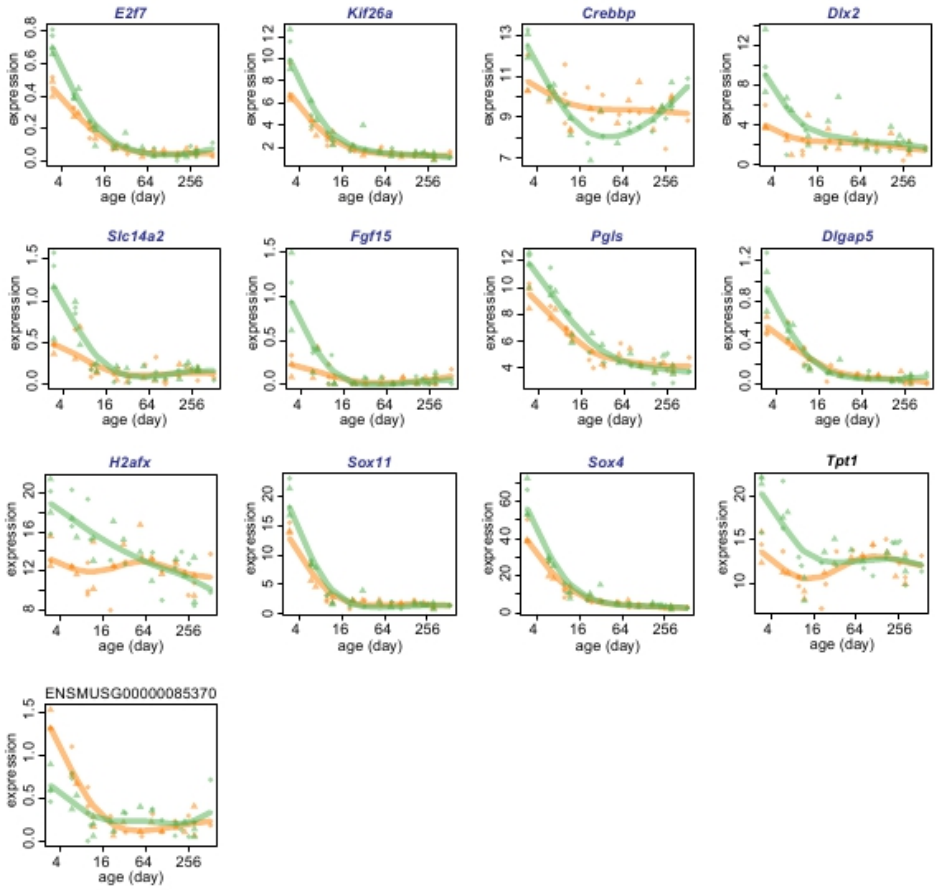

Fig. S9. The expression profiles of the 13 genes with expression affected by the *GLUD2* genotype in the mouse model. Each point represents an individual (orange – transgenic mice; green – control mice). The symbols indicate mice lines: circles – line a; triangles – line b. The lines show spline curves fitted to expression data points with four degrees of freedom. Genes whose names are shown in blue fell into module 1.

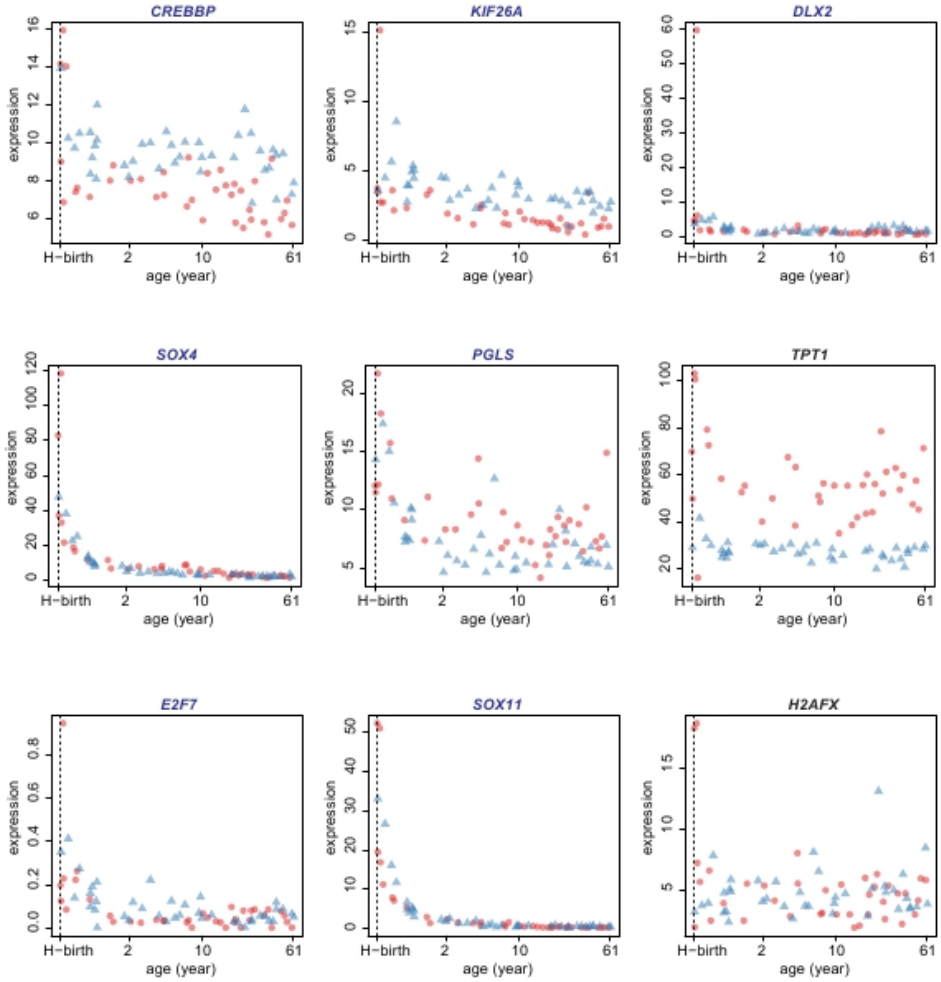

Fig. S10. The expression profiles of the 9 genes with expression affected by the *GLUD2* genotype in our human and macaque dataset. Each point represents an individual. The colors indicate species information: red – humans; blue – macaques. The vertical dashed line shows the human birth age. Genes whose names are shown in blue show the coinciding expression pattern and the coinciding divergence pattern in primates and mice.

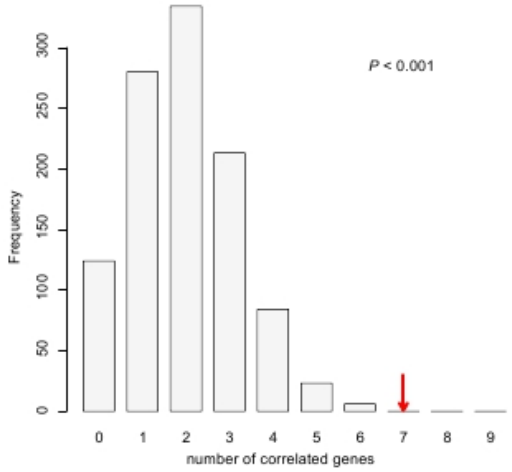

Fig. S11. Frequency of the number of genes showing consistent developmental expression profiles between transgenic mice and humans. The red arrow indicates the number of genes showing consistent expression profiles between transgenic mice and humans among the 9 genes with expression affected by the *GLUD2* genotype in mouse model (Pearson correlation coefficient $r > 0.9$). Bars show the frequency of the number of genes showing consistent expression profiles between transgenic mice and humans calculated by sub-sampling 9 genes from the remaining 23,106 expressed genes in primates, 1,000 times. *P*-value is calculated as the frequency of cases when the number of correlated genes in a sub-sampling is equal or greater than the real value.

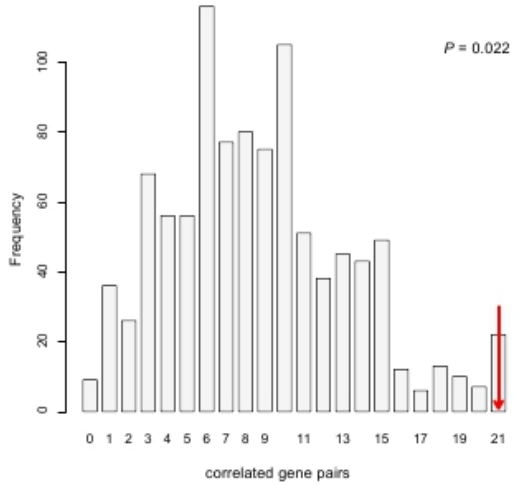

Fig. S12. Frequency of the number of gene pairs showing correlated developmental time-shift between humans and macaques. The red arrow indicates the number of gene pairs showing correlated developmental time-shift between humans and macaques, among the 7 genes showing the coinciding expression and the coinciding divergence pattern in mice and primates (Pearson correlation coefficient $r > 0.9$). Bars show the frequency of the number of gene pairs showing correlated developmental time-shift, calculated by sub-sampling 7 genes from the remaining expressed ones, 1,000 times. *P*-value is calculated as the frequency of cases when the number of correlated gene pairs in a sub-sampling is equal or greater than the real value.

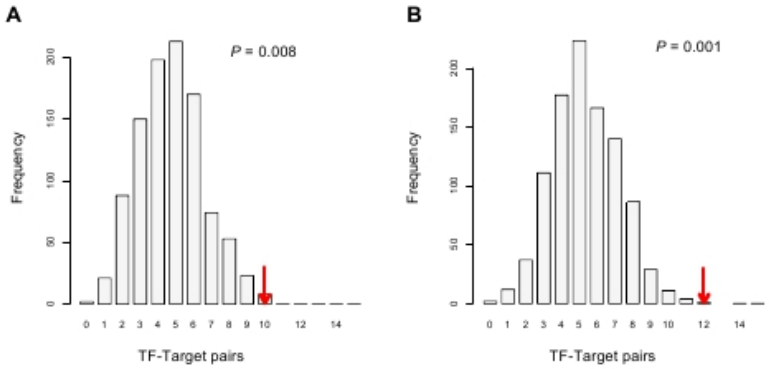

Fig. S13. Frequency of the number of correlated transcription factor – target pairs with regard to divergence. Red arrows show the number of correlated transcription factor – target (TF-Target) pairs with Pearson correlation coefficient calculated based on the divergence of the transcription factors and that of the targets greater than 0.5 ($r >$ 0.5) in mice (A) and primates (B). Bars show the frequency of the number of correlated TF-Target pairs where targets were obtained by sub-sampling the same number of genes as the real target genes, 1,000 times. *P*-value was defined as the frequency of the cases when the number of correlated TF-Target pairs in a sub-sampling is equal or greater than the number of real correlated TF-Target pairs.

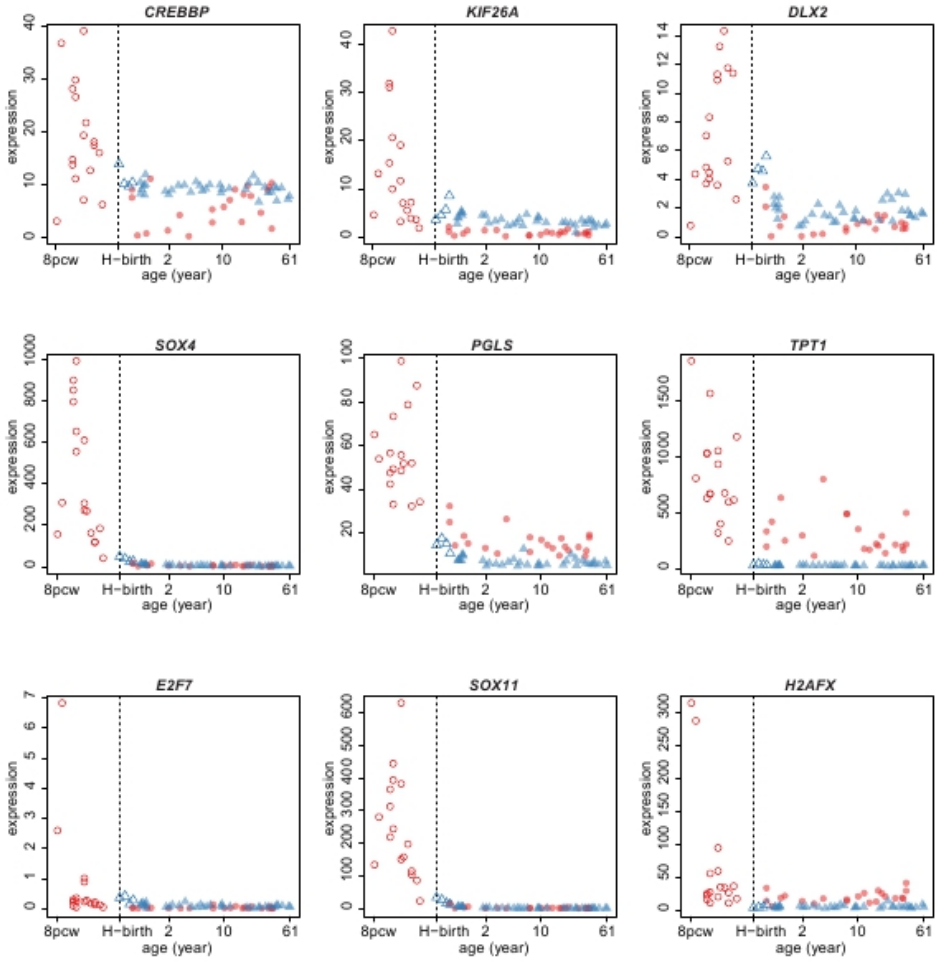

Fig. S14. The expression profiles of the 9 genes with expression affected by the *GLUD2* genotype in the published fetal human and our macaque datasets. Each point represents an individual. The colors indicate species information: red – humans; blue – macaques. Filled symbols indicate postnatal ages, empty symbols - prenatal ages (17 in humans and 4 in macaques). The vertical dashed line shows the human birth age. pcw: weeks post-conception.

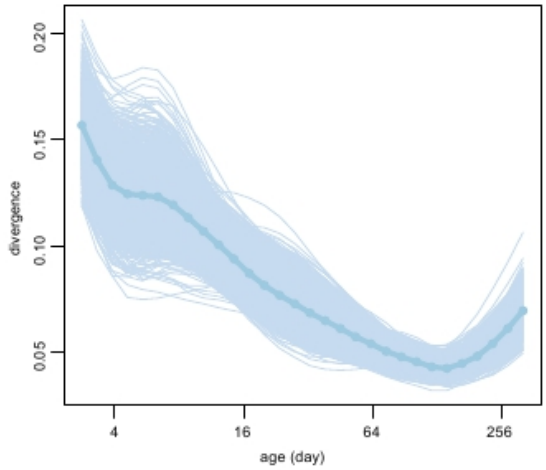

Fig. S15. The metabolite concentration divergence profile of metabolites detected in mouse brain. The mean normalized metabolite concentration divergence between transgenic and control mice based on all the 110 detected metabolites (blue curve). The blue colored area shows the variation of the divergence estimates of the 110 metabolites obtained by bootstrapping 1,000 times.

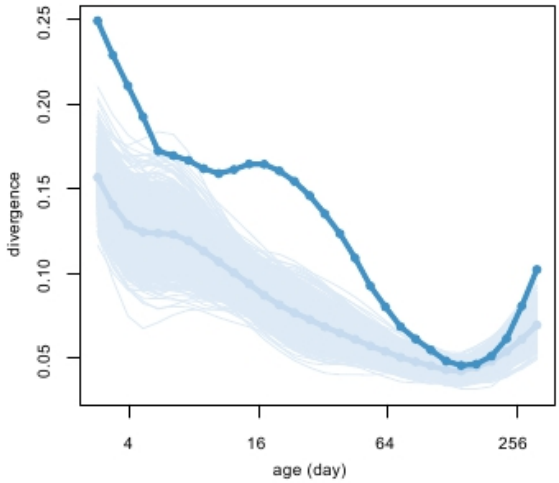

Fig. S16. The metabolite concentration divergence profile in mice, based on the Small Molecule Pathway Database (SMPDB) annotation. The mean normalized metabolite concentration divergence between transgenic and control mice based on the 16 metabolites located in the same SMPDB pathway as the 13 genes with expression affected by the *GLUD2* genotype (dark blue curve) and the remaining 94 metabolites (light blue curve). The light blue colored area shows the variation of divergence estimates of the 94 metabolites, obtained by bootstrapping 1,000 times.

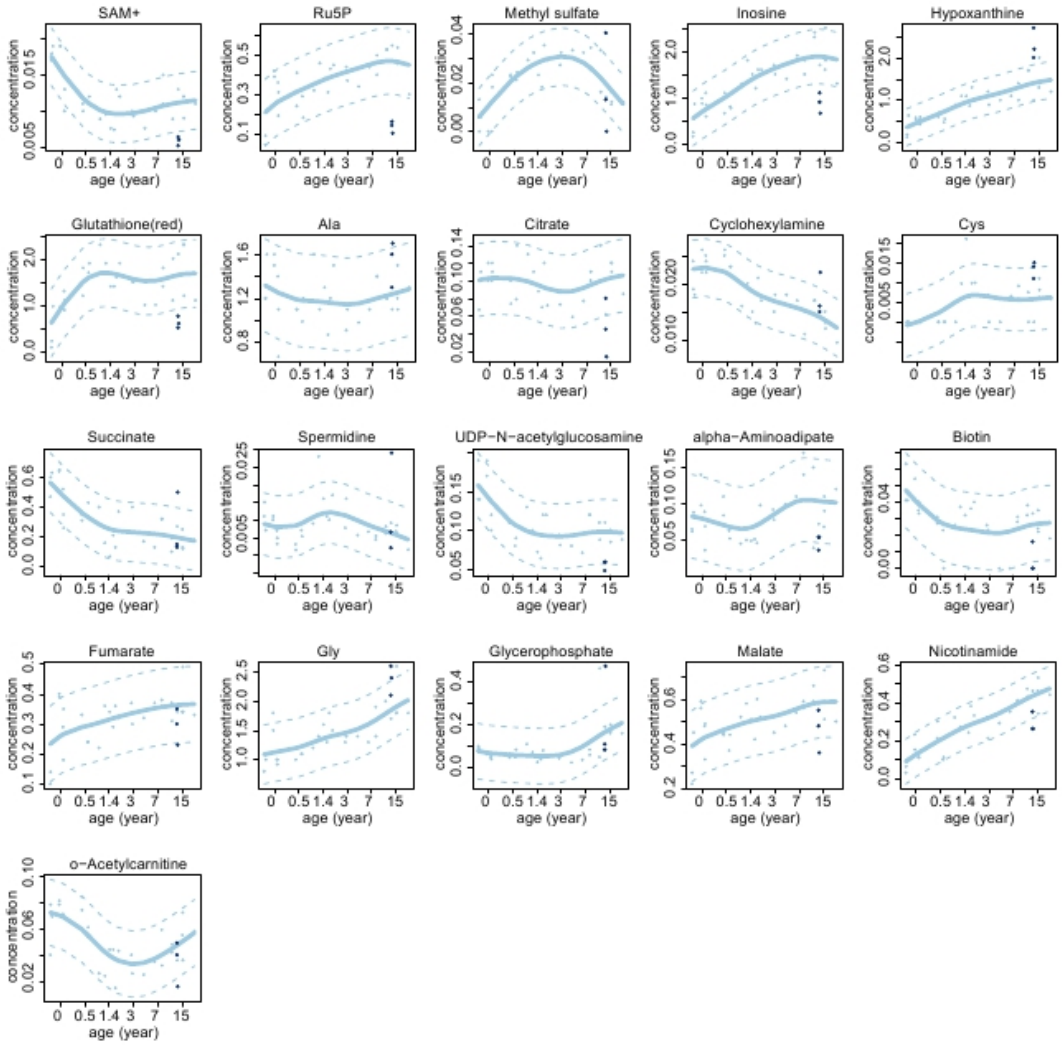

Fig. S17. The concentration profiles of metabolites affected by PMD. The effects of PMD were tested by comparison of the metabolite concentrations of three samples from two macaque individuals (one of them was sampled twice to construct technical replicates) collected with 4-6 hours PMD, with the remaining macaque samples that have a PMD of less than 15 minutes. Light blue points represent macaque samples with short PMD, dark blue points – samples with long PMD. The light blue curves show spline curves fitted to metabolite concentration data points from normal samples with four degrees of freedom. The dashed lines represent the 1.96*standard deviation of the concentration profiles from normal samples.

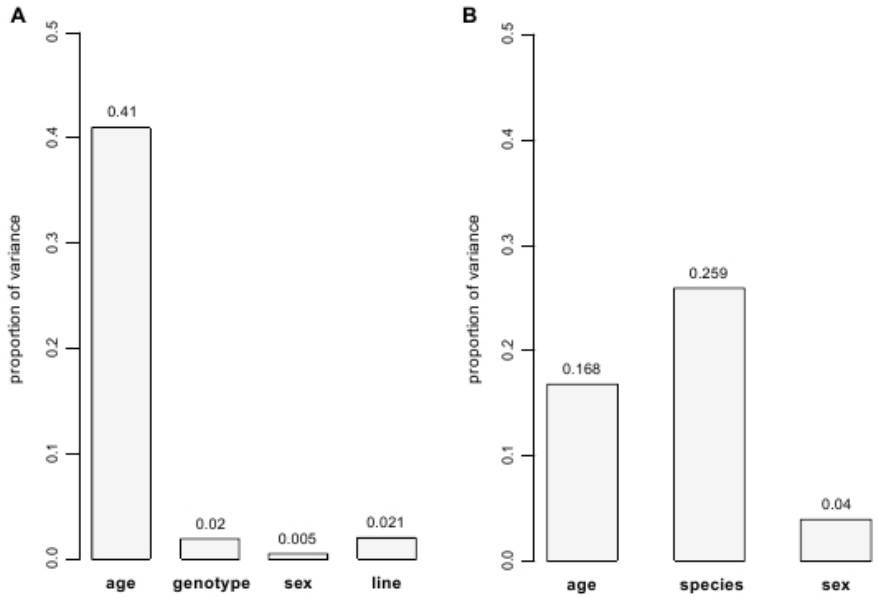

Fig. S18. The proportion of metabolite concentration variation explained by age, genotype/species, sex and line factors. A regression model was fitted to the concentration level of each metabolite of the 110 metabolites detected in mice (A), or the 67 metabolites detected in primates that were not affected by PMD (B), for each of the factors (age, genotype/species, sex and line). The proportion of variance explained by each factor was calculated as one minus the ratio of residual sum of squares from the regression model and residual sum of squares from the null model. The mean proportion of variance explained by each factor across the metabolites is shown above the bars.

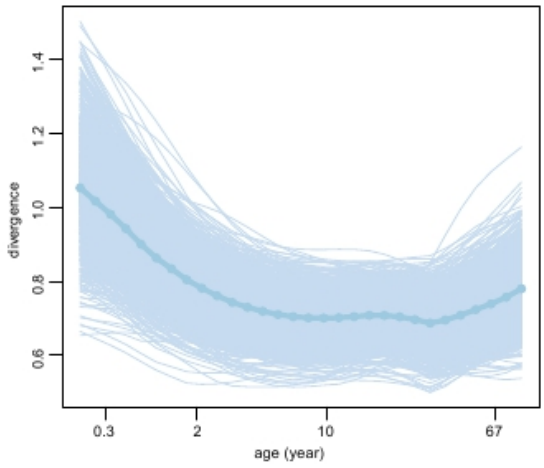

Fig. S19. The metabolite concentration divergence profile of metabolites detected in primate brain. The mean normalized metabolite concentration divergence between humans and macaques based on the 67 detected metabolites that were not affected by PMD (blue curve). The blue colored area shows the variation of divergence estimates of the 67 metabolites obtained by bootstrapping 1,000 times.

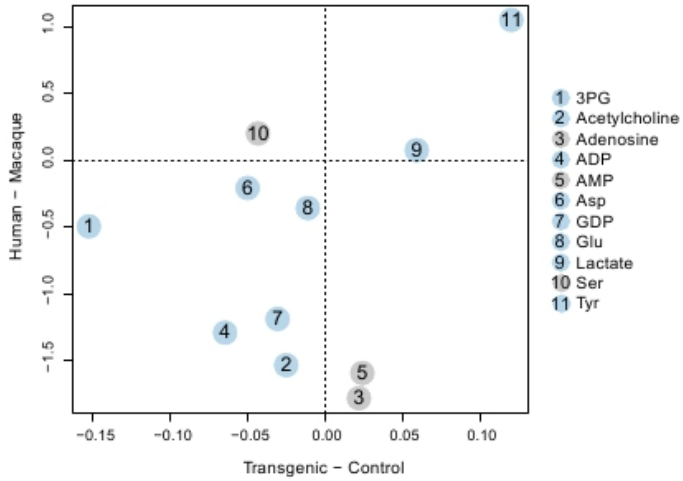

Fig. S20. The direction of metabolite concentration change in primates and transgenic mice. The points show the average normalized metabolite concentration divergence of 11 metabolites linked to *GLUD2* genotype effect in mice and primates measured over the lifespan. Positive values indicate higher concentration level in transgenic mice compared to control mice or higher concentration level in humans compared to macaques. Blue points show metabolites that have consistent direction of divergence in primates and transgenic mice.

SI Tables

Table S1. Mouse sample information

Sample ID	Genotype	Line	Tissue	Gender	Age (Day)	Experiment	RIN	PMD (min)	Weight (mg)
250716	Control	a	CX	M	3	CE-MS;RNA-Seq	10	<5	10.7
250711	Control	a	CX	F	3	CE-MS;RNA-Seq	10	<5	9.8
250712	Control	a	CX	F	3	CE-MS	-	<5	13
460726	Control	b	CX	M	3	CE-MS;RNA-Seq	10	<5	15.3
460728	Control	b	CX	F	3	CE-MS;RNA-Seq	9.8	<5	11.2
460751	Control	b	CX	F	6	CE-MS	-	<5	12
460752	Control	b	CX	F	6	CE-MS	-	<5	11.4
460753	Control	b	CX	F	6	CE-MS	-	<5	11
460754	Control	b	CX	M	6	CE-MS	-	<5	9.4
460755	Control	b	CX	M	6	CE-MS	-	<5	12.7
460756	Control	b	CX	F	6	CE-MS	-	<5	12.4
460757	Control	b	CX	F	6	CE-MS	-	<5	12.7
460758	Control	b	CX	F	6	CE-MS	-	<5	9.4
460760	Control	b	CX	M	6	CE-MS	-	<5	10
460761	Control	b	CX	M	6	CE-MS;RNA-Seq	9.9	<5	12.5
460762	Control	b	CX	M	6	CE-MS	-	<5	10.1
460763	Control	b	CX	M	6	CE-MS	-	<5	11.1
250745	Control	a	CX	F	6	CE-MS	-	<5	10
250746	Control	a	CX	F	6	CE-MS;RNA-Seq	9.9	<5	15.7
250747	Control	a	CX	F	6	CE-MS	-	<5	13.2
250748	Control	a	CX	M	6	CE-MS;RNA-Seq	9.9	<5	11.5
250749	Control	a	CX	M	6	CE-MS	-	<5	11.7
460724	Control	b	CX	M	7	CE-MS	-	<5	11.2
460720	Control	b	CX	F	7	CE-MS	-	<5	9.9
460721	Control	b	CX	F	7	CE-MS;RNA-Seq	9.7	<5	12.5
460722	Control	b	CX	F	7	CE-MS	-	<5	9.7
250735	Control	a	CX	F	10	CE-MS	-	<5	10.6
250737	Control	a	CX	F	10	CE-MS	-	<5	10.4
250739	Control	a	CX	F	10	CE-MS;RNA-Seq	9.6	<5	11.3
250740	Control	a	CX	M	10	CE-MS	-	<5	11.4
250741	Control	a	CX	M	10	CE-MS	-	<5	13.6
250743	Control	a	CX	M	10	CE-MS	-	<5	12.6
250709	Control	a	CX	M	11	CE-MS;RNA-Seq	9.4	<5	15
250710	Control	a	CX	M	11	CE-MS	-	<5	11.5
460710	Control	b	CX	M	12	CE-MS	-	<5	9.7
460711	Control	b	CX	M	12	CE-MS	-	<5	14.8
460740	Control	b	CX	F	12	CE-MS	-	<5	10.7
460741	Control	b	CX	F	12	CE-MS;RNA-Seq	9.4	<5	10.8

460743	Control	b	CX	M	12	CE-MS;RNA-Seq	9.4	<5	9.9
250720	Control	a	CX	F	21	CE-MS	-	<5	9.8
250721	Control	a	CX	F	21	CE-MS	-	<5	10.1
250723	Control	a	CX	M	21	CE-MS	-	<5	11.2
250725	Control	a	CX	M	21	CE-MS	-	<5	10
250726	Control	a	CX	M	21	CE-MS;RNA-Seq	9	<5	12.5
460699	Control	b	CX	F	23	CE-MS;RNA-Seq	9	<5	9.8
460706	Control	b	CX	M	23	CE-MS	-	<5	12.5
250727	Control	a	CX	F	24	CE-MS	-	<5	12.4
460689	Control	b	CX	F	32	CE-MS	-	<5	12.7
460692	Control	b	CX	M	32	CE-MS;RNA-Seq	9.1	<5	9.6
250690	Control	a	CX	F	34	CE-MS;RNA-Seq	9.2	<5	14.1
250700	Control	a	CX	M	34	CE-MS	-	<5	14.9
460672	Control	b	CX	F	55	CE-MS;RNA-Seq	9.2	<5	14.5
460680	Control	b	CX	M	55	CE-MS	-	<5	15.5
250694	Control	a	CX	F	58	CE-MS	-	<5	10.4
250695	Control	a	CX	F	75	CE-MS	-	<5	9.8
460701	Control	b	CX	F	79	CE-MS	-	<5	14.6
460693	Control	b	CX	M	80	CE-MS;RNA-Seq	9.2	<5	9.8
250705	Control	a	CX	M	82	CE-MS;RNA-Seq	9.3	<5	12.2
460670	Control	b	CX	F	111	CE-MS	-	<5	11.5
460678	Control	b	CX	M	111	CE-MS	-	<5	9.8
250685	Control	a	CX	M	111	CE-MS	-	<5	12.7
250684	Control	a	CX	F	111	CE-MS;RNA-Seq	9	<5	12.6
250621	Control	a	CX	M	170	CE-MS;RNA-Seq	9.2	<5	15.4
250618	Control	a	CX	F	170	CE-MS	-	<5	15.5
460644	Control	b	CX	M	176	CE-MS;RNA-Seq	9.2	<5	10.2
460640	Control	b	CX	F	176	CE-MS	-	<5	15.6
460655	Control	b	CX	F	208	CE-MS	-	<5	13
460658	Control	b	CX	M	208	CE-MS	-	<5	14
250597	Control	a	CX	M	210	CE-MS;RNA-Seq	8.9	<5	13.2
250594	Control	a	CX	F	210	CE-MS	-	<5	10.1
460645	Control	b	CX	M	232	CE-MS	-	<5	11.5
460616	Control	b	CX	F	233	CE-MS;RNA-Seq	8.9	<5	9.9
250608	Control	a	CX	F	233	CE-MS	-	<5	12.7
250613	Control	a	CX	M	233	CE-MS	-	<5	9.4
460637	Control	b	CX	M	243	CE-MS	-	<5	15.1
250580	Control	a	CX	F	262	CE-MS	-	<5	14.4
250612	Control	a	CX	M	265	CE-MS;RNA-Seq	8.9	<5	10.5
460613	Control	b	CX	F	266	CE-MS	-	<5	13.5
460589	Control	b	CX	F	316	CE-MS;RNA-Seq	9	<5	10.8
250576	Control	a	CX	F	318	CE-MS;RNA-Seq	8.9	<5	11.1
460579	Control	b	CX	M	319	CE-MS;RNA-Seq	8.9	<5	12.5
250567	Control	a	CX	M	538	RNA-Seq	9	<5	-

250717	Transgenic	a	CX	M	3	CE-MS;RNA-Seq	9.9	<5	13.5
460725	Transgenic	b	CX	M	3	CE-MS;RNA-Seq	10	<5	11.6
460729	Transgenic	b	CX	F	3	CE-MS;RNA-Seq	9.9	<5	12.3
460749	Transgenic	b	CX	F	6	CE-MS	-	<5	13.2
460750	Transgenic	b	CX	F	6	CE-MS	-	<5	13.3
460759	Transgenic	b	CX	F	6	CE-MS;RNA-Seq	10	<5	14.2
250750	Transgenic	a	CX	M	6	CE-MS;RNA-Seq	10	<5	11.2
460723	Transgenic	b	CX	M	7	CE-MS;RNA-Seq	9.8	<5	10.5
250734	Transgenic	a	CX	F	10	CE-MS;RNA-Seq	9.7	<5	9.8
250736	Transgenic	a	CX	F	10	CE-MS	-	<5	10.6
250738	Transgenic	a	CX	F	10	CE-MS;RNA-Seq	9.7	<5	9.9
250742	Transgenic	a	CX	M	10	CE-MS;RNA-Seq	9.8	<5	15.8
460707	Transgenic	b	CX	F	12	CE-MS	-	<5	13.3
460738	Transgenic	b	CX	F	12	CE-MS;RNA-Seq	9.5	<5	9.7
460739	Transgenic	b	CX	F	12	CE-MS	-	<5	10.1
460742	Transgenic	b	CX	M	12	CE-MS;RNA-Seq	9.6	<5	10.2
250724	Transgenic	a	CX	M	21	CE-MS;RNA-Seq	9	<5	13.2
460700	Transgenic	b	CX	F	23	CE-MS;RNA-Seq	9	<5	11.5
460705	Transgenic	b	CX	M	23	CE-MS	-	<5	9.8
250728	Transgenic	a	CX	F	24	CE-MS;RNA-Seq	8.8	<5	11.8
460690	Transgenic	b	CX	F	32	CE-MS;RNA-Seq	good	<5	11
460691	Transgenic	b	CX	M	32	CE-MS	-	<5	10
250691	Transgenic	a	CX	F	34	CE-MS	-	<5	13.6
250701	Transgenic	a	CX	M	34	CE-MS;RNA-Seq	8.7	<5	13.3
460675	Transgenic	b	CX	F	55	CE-MS;RNA-Seq	8.8	<5	14.7
460679	Transgenic	b	CX	M	55	CE-MS	-	<5	10.6
250693	Transgenic	a	CX	F	58	CE-MS	-	<5	11.5
250696	Transgenic	a	CX	M	58	CE-MS;RNA-Seq	8.6	<5	9.6
250697	Transgenic	a	CX	M	58	CE-MS	-	<5	13.3
250692	Transgenic	a	CX	F	75	CE-MS	-	<5	14
460702	Transgenic	b	CX	F	79	CE-MS;RNA-Seq	8.5	<5	15.7
460694	Transgenic	b	CX	M	80	CE-MS	-	<5	10.8
250706	Transgenic	a	CX	M	82	CE-MS;RNA-Seq	8.7	<5	11.4
460671	Transgenic	b	CX	F	111	CE-MS	-	<5	13.9
460677	Transgenic	b	CX	M	111	CE-MS;RNA-Seq	8.6	<5	14.7
250686	Transgenic	a	CX	M	111	CE-MS	-	<5	11.1
250683	Transgenic	a	CX	F	111	CE-MS;RNA-Seq	8.6	<5	12.6
250622	Transgenic	a	CX	M	170	CE-MS;RNA-Seq	8.8	<5	10.8
250619	Transgenic	a	CX	F	170	CE-MS	-	<5	11.5
460643	Transgenic	b	CX	M	176	CE-MS	-	<5	10
460639	Transgenic	b	CX	F	176	CE-MS;RNA-Seq	8.6	<5	15.6
460654	Transgenic	b	CX	F	208	CE-MS	-	<5	11.8
460657	Transgenic	b	CX	M	208	CE-MS	-	<5	12.7
250599	Transgenic	a	CX	M	210	CE-MS	-	<5	9.9

250595	Transgenic	a	CX	F	210	CE-MS	-	<5	13.7
460642	Transgenic	b	CX	M	232	CE-MS;RNA-Seq	8.7	<5	10.6
460615	Transgenic	b	CX	F	233	CE-MS	-	<5	10
250607	Transgenic	a	CX	F	233	CE-MS	-	<5	10.8
250610	Transgenic	a	CX	M	233	CE-MS	-	<5	11.5
460638	Transgenic	b	CX	M	243	CE-MS	-	<5	10.4
250581	Transgenic	a	CX	F	262	CE-MS;RNA-Seq	8.5	<5	13.8
250611	Transgenic	a	CX	M	265	CE-MS	-	<5	15.4
460612	Transgenic	b	CX	F	266	CE-MS	-	<5	9.6
460588	Transgenic	b	CX	F	316	CE-MS;RNA-Seq	8.6	<5	11
250577	Transgenic	a	CX	F	318	CE-MS;RNA-Seq	8.6	<5	10.9
460581	Transgenic	b	CX	M	319	CE-MS;RNA-Seq	8.6	<5	11.5
250566	Transgenic	a	CX	M	538	RNA-Seq	8.6	<5	-

CX: frontal cortical region

RIN: RNA integrity values for RNA-Seq experiment

PMD: postmortem delay duration

Weight: sample weight (mg) for CE-MS experiment

Table S2. Primate sample information

Sample ID	Species	Tissue	Sex	Age		Experiment	RIN	PMD (h)	Weight
				year	day				
447	Human	PFC	M	0	2	RNA-seq	8	3	-
779	Human	PFC	M	0	5	RNA-seq	8.8	5	-
398	Human	PFC	F	0	16	CE-MS; RNA-seq	8.8	3	9.7
1157	Human	PFC	F	0	20	CE-MS; RNA-seq	7.1	14	10.7
2008267	Human	PFC	F	0	30	CE-MS	-	-	15.5
759	Human	PFC	M	0	35	CE-MS; RNA-seq	7.9	7	10
1055	Human	PFC	M	0	96	RNA-seq	7.7	12	-
5183	Human	PFC	M	0	107	RNA-seq	7.7	-	-
1325	Human	PFC	F	0	182	CE-MS	-	1	12.6
131	Human	PFC	F	0	198	CE-MS	-	24	10.4
1303	Human	PFC	M	0	212	RNA-seq	8	8	-
1453	Human	PFC	M	1	78	RNA-seq	6.7	19	-
814	Human	PFC	M	1	123	RNA-seq	7.6	19	-
1063	Human	PFC	M	1	123	CE-MS	-	21	11.2
1488	Human	PFC	M	1	137	CE-MS	-	21	12.7
1275	Human	PFC	F	2	57	RNA-seq	5.5	21	-
510	Human	PFC	F	2	171	CE-MS	-	20	11.5
1791	Human	PFC	M	2	286	CE-MS; RNA-seq	8.2	12	13.7
6736	Human	PFC	F	4	0	RNA-seq	6.8	-	-
1185	Human	PFC	M	4	258	CE-MS; RNA-seq	8.6	17	11.7
4907	Human	PFC	F	4	274	CE-MS; RNA-seq	8.1	15	17.7
4898	Human	PFC	M	7	272	RNA-seq	6.4	12	-
629	Human	PFC	M	7	306	CE-MS	-	18	12.2
1860	Human	PFC	M	8	2	CE-MS; RNA-seq	8	5	10.2
1706	Human	PFC	F	8	214	RNA-seq	7.1	20	-
2011296	Human	PFC	M	9	0	CE-MS	-	-	13.7
5161	Human	PFC	F	10	262	CE-MS; RNA-seq	7.4	22	9.4
M3228	Human	PFC	M	11	294	RNA-seq	8.3	-	-
1908	Human	PFC	M	13	360	CE-MS; RNA-seq	8.2	13	14
1024	Human	PFC	M	14	60	CE-MS	-	16	10
5242	Human	PFC	M	15	119	RNA-seq	8.4	-	-
4848	Human	PFC	M	16	271	CE-MS	-	15	11.3
7387	Human	PFC	M	17	0	RNA-seq	8.1	-	-
2009325	Human	PFC	M	17	0	CE-MS	-	-	9.5
1571	Human	PFC	F	18	138	CE-MS	-	8	15.5
5251	Human	PFC	M	19	0	RNA-seq	6.6	-	-
1011	Human	PFC	F	19	69	CE-MS	-	7	11.2
4548	Human	PFC	F	20	63	RNA-seq	8.7	5	-
1846	Human	PFC	F	20	221	CE-MS; RNA-seq	7.1	9	14
933	Human	PFC	M	20	255	CE-MS	-	12	12.1

1442	Human	PFC	M	22	322	RNA-seq	7.5	7	-
7738	Human	PFC	M	24	0	RNA-seq	7.2	-	-
1455	Human	PFC	F	25	149	CE-MS	-	7	10
602	Human	PFC	M	27	42	CE-MS; RNA-seq	8.8	15	15.1
B-24	Human	PFC	F	28	0	RNA-seq	1.5 (ok)	24	-
1502	Human	PFC	M	29	363	RNA-seq	7.4	19	-
7344	Human	PFC	F	36	0	RNA-seq	6.6	-	-
7561	Human	PFC	M	39	0	RNA-seq	7	-	-
1134	Human	PFC	M	41	241	CE-MS; RNA-seq	8.9	15	10.8
4841	Human	PFC	F	42	127	CE-MS	-	17	12.8
6259	Human	PFC	M	50	0	RNA-seq	8	-	-
1474	Human	PFC	M	50	156	CE-MS	-	8	15.8
5117	Human	PFC	M	50	297	CE-MS	-	5	11.8
1578	Human	PFC	M	53	112	RNA-seq	8.3	-	-
6860	Human	PFC	M	56	0	RNA-seq	8	6	-
S03_074	Human	PFC	F	61	0	CE-MS	-	6.83	12
4263	Human	PFC	M	61	187	CE-MS; RNA-seq	8.8	6	12.8
S01_118	Human	PFC	M	88	0	CE-MS	-	7.42	10.5
5089	Human	PFC	M	89	18	CE-MS	-	14	10.5
S96_297	Human	PFC	F	90	0	CE-MS	-	6.17	14.5
f0507910	Macaque	PFC	F	0	-70	RNA-seq	9.3	<0.3	-
f9604682	Macaque	PFC	M	0	-56	RNA-seq	8.6	<0.3	-
F0909	Macaque	PFC	M	0	-42	RNA-seq	9	<0.3	-
F0908	Macaque	PFC	M	0	-36	CE-MS	-	<0.3	9.8
F0910	Macaque	PFC	M	0	-35	CE-MS	-	<0.3	11
F0906	Macaque	PFC	M	0	-30	CE-MS; RNA-seq	9.1	<0.3	13.8
NB0903	Macaque	PFC	M	0	0.5	RNA-seq	OK	<0.3	-
NB0901	Macaque	PFC	M	0	1	CE-MS; RNA-seq	8.5	<0.3	10.9
NB0904	Macaque	PFC	F	0	2	CE-MS	-	<0.3	10.2
NB0905	Macaque	PFC	M	0	7	RNA-seq	9.3	<0.3	-
705	Macaque	PFC	M	0	16	RNA-seq	9.1	<0.3	-
704	Macaque	PFC	M	0	20	CE-MS; RNA-seq	OK	<0.3	10.1
703	Macaque	PFC	M	0	22	CE-MS	-	<0.3	9.8
702	Macaque	PFC	M	0	23	CE-MS; RNA-seq	9.5	<0.3	11.5
701	Macaque	PFC	M	0	24	RNA-seq	9	<0.3	-
70175	Macaque	PFC	M	0	151	CE-MS; RNA-seq	9.6	<0.3	9.8
70141	Macaque	PFC	M	0	153	CE-MS	-	<0.3	10.2
70115	Macaque	PFC	M	0	179	RNA-seq	OK	<0.3	-
70133	Macaque	PFC	M	0	207	CE-MS; RNA-seq	OK	<0.3	10.8
6403	Macaque	PFC	M	0	215	CE-MS	-	<0.3	10.2
6709	Macaque	PFC	M	0	278	RNA-seq	OK	<0.3	-
6237	Macaque	PFC	M	0	353	CE-MS; RNA-seq	9	<0.3	10.7
61569	Macaque	PFC	M	1	80	CE-MS; RNA-seq	8.3	<0.3	9.6

51087	Macaque	PFC	M	1	170	CE-MS; RNA-seq	9.3	<0.3	12.1
51373	Macaque	PFC	M	1	242	CE-MS; RNA-seq	9.1	<0.3	15
51469	Macaque	PFC	M	1	294	CE-MS; RNA-seq	OK	<0.3	9.9
51095	Macaque	PFC	M	2	9	CE-MS; RNA-seq	9	<0.3	9.9
5773	Macaque	PFC	M	2	101	RNA-seq	OK	<0.3	-
4093	Macaque	PFC	M	3	40	CE-MS; RNA-seq	9.3	<0.3	12.6
50715	Macaque	PFC	M	3	80	CE-MS; RNA-seq	OK	<0.3	10
4089	Macaque	PFC	M	3	110	RNA-seq	OK	<0.3	-
3071	Macaque	PFC	M	4	27	RNA-seq	7.8	<0.3	-
m0507910	Macaque	PFC	F	4	227	CE-MS	-	<0.3	10
B00051	Macaque	PFC	M	6	165	CE-MS; RNA-seq	OK	<0.3	11
135	Macaque	PFC	M	7	15	RNA-seq	8.7	<0.3	-
99057	Macaque	PFC	M	8	16	CE-MS; RNA-seq	9.3	<0.3	12.4
MM0906	Macaque	PFC	F	8	34	CE-MS	-	<0.3	10.1
98145	Macaque	PFC	M	9	37	RNA-seq	7.8	<0.3	-
98073	Macaque	PFC	M	9	104	RNA-seq	8.5	<0.3	-
5917	Macaque	PFC	M	10	291	CE-MS	-	<0.3	9.6
96007	Macaque	PFC	M	10	328	CE-MS; RNA-seq	8.4	<0.3	10.3
95001	Macaque	PFC	M	11	346	RNA-seq	8.4	<0.3	-
m9704920	Macaque	PFC	F	12	263	CE-MS	-	<0.3	11.9
94051	Macaque	PFC	M	13	17	RNA-seq	8.7	<0.3	-
93041	Macaque	PFC	M	14	21	RNA-seq	OK	<0.3	-
9605799	Macaque	PFC	M	14	297	CE-MS	-	<0.3	11.9
92095	Macaque	PFC	M	14	349	CE-MS; RNA-seq	OK	<0.3	11.9
92107	Macaque	PFC	M	15	3	CE-MS; RNA-seq	OK	<0.3	13.2
90049	Macaque	PFC	M	17	22	RNA-seq	9	<0.3	-
87015	Macaque	PFC	M	20	91	RNA-seq	OK	<0.3	-
86023	Macaque	PFC	M	21	8	CE-MS; RNA-seq	OK	<0.3	13.8

PFC: prefrontal cortex

RIN: RNA integrity values for RNA-Seq experiment

PMD: postmortem delay duration

Weight: sample weight (mg) for CE-MS experiment

Table S3. Numbers of sequenced reads and mapped reads in mice

Sample ID	Genotype	Line	Tissue	Age (Day)	Total reads	Mapped reads	Proportion
250716	Control	a	CX	3	12309907	10436865	84.78%
250711	Control	a	CX	3	11443680	9654611	84.37%
460726	Control	b	CX	3	13165498	11158041	84.75%
460728	Control	b	CX	3	14376674	12095430	84.13%
460761	Control	b	CX	6	14101841	11945518	84.71%
250746	Control	a	CX	6	13294674	11223830	84.42%
250748	Control	a	CX	6	12587762	10635109	84.49%
460721	Control	b	CX	7	13569840	11478775	84.59%
250739	Control	a	CX	10	11837570	10020922	84.65%
250709	Control	a	CX	11	10833822	9197671	84.90%
460741	Control	b	CX	12	11701648	9911836	84.70%
460743	Control	b	CX	12	13087909	11110004	84.89%
250726	Control	a	CX	21	11889249	10055126	84.57%
460699	Control	b	CX	23	12135829	10244495	84.42%
460692	Control	b	CX	32	13905382	11721231	84.29%
250690	Control	a	CX	34	13996250	11861518	84.75%
460672	Control	b	CX	55	13154111	11093174	84.33%
460693	Control	b	CX	80	11271887	9480569	84.11%
250705	Control	a	CX	82	14190375	11966085	84.33%
250684	Control	a	CX	111	13413469	11398975	84.98%
250621	Control	a	CX	170	10663251	9018811	84.58%
460644	Control	b	CX	176	12894776	10990189	85.23%
250597	Control	a	CX	210	14933485	12714867	85.14%
460616	Control	b	CX	233	11904159	10086876	84.73%
250612	Control	a	CX	265	13298212	11233825	84.48%
460589	Control	b	CX	316	13523261	11486244	84.94%
250576	Control	a	CX	318	12014453	10243683	85.26%
460579	Control	b	CX	319	14025055	11974878	85.38%
250567	Control	a	CX	538	11896917	10090660	84.82%
250717	Transgenic	a	CX	3	10951151	9265528	84.61%
460725	Transgenic	b	CX	3	11938746	10115603	84.73%
460729	Transgenic	b	CX	3	11596396	9778462	84.32%
460759	Transgenic	b	CX	6	13613851	11503131	84.50%
250750	Transgenic	a	CX	6	12549257	10558232	84.13%
460723	Transgenic	b	CX	7	10744051	9102892	84.72%
250734	Transgenic	a	CX	10	10330716	8845545	85.62%
250738	Transgenic	a	CX	10	11272990	9623665	85.37%
250742	Transgenic	a	CX	10	12134776	10201207	84.07%
460738	Transgenic	b	CX	12	11832083	9924828	83.88%
460742	Transgenic	b	CX	12	11553719	9843701	85.20%

250724	Transgenic	a	CX	21	11319072	9622155	85.01%
460700	Transgenic	b	CX	23	11023624	9348739	84.81%
250728	Transgenic	a	CX	24	10949231	9295508	84.90%
460690	Transgenic	b	CX	32	12193304	10314568	84.59%
250701	Transgenic	a	CX	34	13208091	11153216	84.44%
460675	Transgenic	b	CX	55	11918819	10069035	84.48%
250696	Transgenic	a	CX	58	12292848	10389521	84.52%
460702	Transgenic	b	CX	79	12050587	10175172	84.44%
250706	Transgenic	a	CX	82	12615602	10631615	84.27%
460677	Transgenic	b	CX	111	12174180	10322281	84.79%
250683	Transgenic	a	CX	111	12363646	10450296	84.52%
250622	Transgenic	a	CX	170	11102862	9420050	84.84%
460639	Transgenic	b	CX	176	12875879	10913488	84.76%
460642	Transgenic	b	CX	232	12524168	10605826	84.68%
250581	Transgenic	a	CX	262	11134385	9445391	84.83%
460588	Transgenic	b	CX	316	11248567	9590992	85.26%
250577	Transgenic	a	CX	318	12149693	10349800	85.19%
460581	Transgenic	b	CX	319	10483007	8925913	85.15%
250566	Transgenic	a	CX	538	10734205	9104225	84.82%

Total reads: total number of raw reads generated by RNA-Seq

Mapped reads: uniquely mapped reads using mapping program “STAR”

Proportion: the ratio of uniquely mapped reads and total reads

Table S4. Tags from both strands of the *GLUD2* promoter region obtained from deepCAGE data from FANTOM3 and FANTOM4

chromosome	coordinate	strand	Number of tags
chrX	120009065	+	1
chrX	120009146	+	8
chrX	120009206	-	3
chrX	120009308	-	1
chrX	120009501	-	1
chrX	120009661	+	1

Coordinate: the coordinate is based on human reference genome hg18.

Table S5. 13 genes differentially expressed between transgenic and control mice

Ensembl ID	Module	HGNC ID	Description	Expressed in primates
ENSMUSG00000020185	1	*E2F7*	E2F transcription factor 7	Yes
ENSMUSG00000022521	1	*Crebbp*	CREB binding protein	Yes
ENSMUSG00000023391	1	*Dlx2*	Distal-less homeobox 2	Yes
ENSMUSG00000024552	1	*Slc14a2*	Solute carrier family 14 (urea transporter), member 2	No
ENSMUSG00000031073	1	*Fgf15*	Fibroblast growth factor 15	No
ENSMUSG00000031807	1	*Pgls*	6-phosphogluconolactonase	Yes
ENSMUSG00000037544	1	*Dlgap5*	discs, large (Drosophila) homolog-associated protein 5	No
ENSMUSG00000049932	1	*H2afx*	H2A histone family, member X	Yes
ENSMUSG00000063632	1	*Sox11*	SRY (sex determining region Y)-box 11	Yes
ENSMUSG00000076431	1	*Sox4*	SRY (sex determining region Y)-box 4	Yes
ENSMUSG00000021294	1	*Kif26a*	kinesin family member 26A	Yes
ENSMUSG00000060126	4	*Tpt1*	Tumor protein, translationally-controlled 1	Yes
ENSMUSG00000085370	4	2310002F09Rik	RIKEN cDNA 2310002F09 gene	No

Module: the co-expression module the gene fell into in the unsupervised hierarchical clustering
Expressed in primates: indicates whether the gene was expressed in our human and rhesus macaque dataset

Table S6. Functional enrichment analysis of 11 co-expressed genes affected by the *GLUD2* genotype in the mouse model using "GeneCodis"

GO biological process	Support	Size	Reference Support	Reference Size	*p*-value
noradrenergic neuron differentiation	2	11	2	15661	2.76E-05
neuroepithelial cell differentiation	2	11	2	15661	2.76E-05
glial cell development	2	11	3	15661	4.14E-05
cardiac ventricle formation	2	11	3	15661	4.14E-05
glial cell proliferation	2	11	4	15661	6.62E-05
sympathetic nervous system development	2	11	6	15661	1.38E-04
limb bud formation	2	11	7	15661	1.65E-04
spinal cord development	2	11	8	15661	1.93E-04
neural tube formation	2	11	14	15661	5.55E-04
ventricular septum morphogenesis	2	11	15	15661	5.76E-04
negative regulation of cell death	2	11	27	15661	1.74E-03
positive regulation of transcription from RNA polymerase II promoter	4	11	459	15661	2.09E-03
skeletal system development	2	11	56	15661	5.94E-03
positive regulation of cell proliferation	3	11	258	15661	6.26E-03
positive regulation of transcription, DNA-dependent	3	11	389	15661	7.19E-03
negative regulation of transcription from RNA polymerase II promoter	3	11	370	15661	7.95E-03
regulation of transcription, DNA-dependent	5	11	1339	15661	8.31E-03
heart development	2	11	122	15661	8.64E-03
cell cycle	3	11	487	15661	1.01E-02
transcription, DNA-dependent	4	11	1254	15661	1.59E-02
negative regulation of cell proliferation	2	11	219	15661	1.76E-02
cell differentiation	2	11	403	15661	3.87E-02

Support: number of annotated genes among the co-expressed genes

Size: total number of co-expressed genes

Reference support: number of annotated genes among all the age-dependent genes

Reference size: total number of age-dependent genes

p-value: hypergeometric test *p*-value after FDR correction

Table S7. Target genes of *SOX4*, *SOX11*, *DLX2*

TF	Target			Reference	Primate
	Ensembl ID	HGNC	Description		
DLX2	ENSMUSG00000055254	*Ntrk2*	Neurotrophic tyrosine kinase, receptor, type2	(8)	Yes
	ENSMUSG00000035277	*Arx*	Aristaless related homeobox	(9)	Yes
	ENSMUSG00000021994	*Wnt5a*	Wingless-type MMTV integration site family, member 5A	(10)	Yes
	ENSMUSG00000015812	*Gnrh1*	Gonadotropin releasing hormone 1	(11)	Yes
	ENSMUSG00000025969	*Nrp2*	neuropilin 2	(12)	Yes
SOX11	ENSMUSG00000015812	*Gnrh1*	Gonadotropin releasing hormone 1	(13)	Yes
	ENSMUSG00000004891	*Nes*	Nestin	(14)	Yes
	ENSMUSG00000028284	*Map3k7*	Mitogen-activated protein kinase kinase kinase 7	(15)	Yes
	ENSMUSG00000020745	*Pafah1b1*	Platelet-activating factor acetylhydrolase, isoform 1b, subunit 1	(15)	Yes
	ENSMUSG00000037169	*Mycn*	V-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian)	(15)	Yes
	ENSMUSG00000031285	*Dcx*	Doublecortin	(16)	Yes
	ENSMUSG00000021743	*Fezf2*	Fez family zinc finger 2	(17)	Yes
	ENSMUSG00000064289	*Tank*	TRAF family member-associated Nf-kappa B activator	(18)	Yes
	ENSMUSG00000019817	*Plagl1*	pleiomorphic adenoma gene-like 1	(19)	Yes
	ENSMUSG00000062380	*Tubb3*	Tubulin, beta 3 class III	(20)	No
SOX4	ENSMUSG00000030796	*Tead2*	TEA domain family member 2	(21)	Yes
	ENSMUSG00000021743	*Fezf2*	Fez family zinc finger 2	(17)	Yes
	ENSMUSG00000062380	*Tubb3*	Tubulin, beta 3 class III	(20)	No

Primate: indicates whether the gene was expressed in human and rhesus macaque dataset

Table S8. Metabolites detected in the mouse model

KEGG ID	Compound name	Pathway sharing
C00037	Gly	Yes
C00041	Ala	Yes
C00065	Ser	Yes
C00049	Asp	Yes
C01996	Acetylcholine	Yes
C00025	Glu	Yes
C00082	Tyr	Yes
C00212	Adenosine	Yes
C00186	Lactate	Yes
C00122	Fumarate	Yes
C00042	Succinate	Yes
C00074	PEP	Yes
C00197	3PG	Yes
C00158	Citrate	Yes
C00199	Ru5P	Yes
C00085	F6P	Yes
C00345	6-Phosphogluconate	Yes
C05382	S7P	Yes
C00354	F1,6P	Yes
C00020	AMP	Yes
C00008	ADP	Yes
C00035	GDP	Yes
C00002	ATP	Yes
C00044	GTP	Yes
C00086	Urea	No
C01104	Trimethylamine N-oxide	No
C00099	beta-Ala	No
C00571	Cyclohexylamine	No
C02356	2AB	No
C00334	GABA	No
C00114	Choline	No
C00519	Hypotaurine	No
C00386	Carnosine	No
C00791	Creatinine	No
C00148	Pro	No
C00581	Guanidinoacetate	No
C00183	Val	No
C00719	Betaine	No
C00188	Thr	No
C00263	Homoserine	No
C00153	Nicotinamide	No

C05127	1-Methylhistamine	No
C00408	Pipecolate	No
C01015	Hydroxyproline	No
C00300	Creatine	No
C00407	Ile	No
C00123	Leu	No
C00152	Asn	No
C00077	Ornithine	No
C00147	Adenine	No
C00262	Hypoxanthine	No
C02918	1-Methylnicotinamide	No
C01035	gamma-Guanidinobutyrate	No
C01181	gamma-Butyrobetaine	No
C00315	Spermidine	No
C00064	Gln	No
C00047	Lys	No
C00073	Met	No
C00135	His	No
C00956	alpha-Aminoadipate	No
C00318	Carnitine	No
C00079	Phe	No
C01152	3-Methylhistidine	No
C00062	Arg	No
C00327	Citrulline	No
C00588	Phosphorylcholine	No
C03793	N6,N6,N6-Trimethyllysine	No
C00021	SAH	No
C00019	SAM+	No
C02571	o-Acetylcarnitine	No
C00078	Trp	No
C00884	Homocarnosine	No
C00475	Cytidine	No
C00647	Pyridoxamine 5'-phosphate	No
C00670	Glycerophosphorylcholine	No
C00294	Inosine	No
C02494	1-Methyladenosine	No
C00387	Guanosine	No
C00170	5-Methylthioadenosine	No
C00127	Glutathione(ox)	No
C00051	Glutathione(red)	No
C01089	3-Hydroxybutyrate	No
C02704	Methyl sulfate	No
C05123	Isethionate	No
C01879	5-Oxoproline	No

C00711	Malate	No
C00346	Ethanolamine phosphate	No
C02630	2-Hydroxyglutarate	No
C00423	trans-Cinnamate	No
C03761	3-Hydroxy-3-methylglutarate	No
C00093	Glycerophosphate	No
C00417	cis-Aconitate	No
C01042	N-Acetylaspartate	No
C00624	N-Acetylglutamate	No
C00864	Pantothenate	No
C00092	G6P	No
C00103	G1P	No
C04501	N-Acetylglucosamine 1-phosphate	No
C00270	N-Acetylneuraminate	No
C00055	CMP	No
C00105	UMP	No
C00130	IMP	No
C00144	GMP	No
C00112	CDP	No
C00015	UDP	No
C00689	Trehalose 6-phosphate	No
C03794	Adenylosuccinate	No
C00029	UDP-glucose	No
C00167	UDP-glucuronate	No
C00043	UDP-N-acetylglucosamine	No

Pathway sharing: indicates whether the metabolite was in the same KEGG pathway as the 13 differentially expressed genes affected by *GLUD2* genotype in the mouse model

Table S9. Metabolites detected in humans and rhesus macaques

KEGG ID	Compound name	Affected by PMD	Linked from mice
C00197	3PG	No	Yes
C01996	Acetylcholine	No	Yes
C00212	Adenosine	No	Yes
C00008	ADP	No	Yes
C00020	AMP	No	Yes
C00049	Asp	No	Yes
C00035	GDP	No	Yes
C00025	Glu	No	Yes
C00186	Lactate	No	Yes
C00065	Ser	No	Yes
C00082	Tyr	No	Yes
C02918	1-Methylnicotinamide	No	No
C02356	2AB	No	No
C03761	3-Hydroxy-3-methylglutarate	No	No
C01089	3-Hydroxybutyrate	No	No
C01152	3-Methylhistidine	No	No
C01879	5-Oxoproline	No	No
C00147	Adenine	No	No
C00062	Arg	No	No
C00152	Asn	No	No
C00099	beta-Ala	No	No
C00719	Betaine	No	No
C00318	Carnitine	No	No
C00114	Choline	No	No
C00327	Citrulline	No	No
C00300	Creatine	No	No
C00791	Creatinine	No	No
C02291	Cystathionine	No	No
C00475	Cytidine	No	No
C00346	Ethanolamine phosphate	No	No
C00092	G6P	No	No
C00334	GABA	No	No
C01181	gamma-Butyrobetaine	No	No
C01035	gamma-Guanidinobutyrate	No	No
C00064	Gln	No	No
C00127	Glutathione(ox)	No	No
C00670	Glycerophosphorylcholine	No	No
C00144	GMP	No	No
C00387	Guanosine	No	No
C00135	His	No	No
C00884	Homocarnosine	No	No

C01015	Hydroxyproline	No	No
C00519	Hypotaurine	No	No
C00407	Ile	No	No
C00130	IMP	No	No
C00123	Leu	No	No
C00047	Lys	No	No
C00073	Met	No	No
C01042	N-Acetylaspartate	No	No
C00624	N-Acetylglutamate	No	No
C02997	N-Acetylhistidine	No	No
C00270	N-Acetylneuraminate	No	No
C03793	N6,N6,N6-Trimethyllysine	No	No
C00077	Ornithine	No	No
C00864	Pantothenate	No	No
C00079	Phe	No	No
C00588	Phosphorylcholine	No	No
C00148	Pro	No	No
C00134	Putrescine(1,4-Butanediamine)	No	No
C00021	SAH	No	No
C00188	Thr	No	No
C00423	trans-Cinnamate	No	No
C00689	Trehalose 6-phosphate	No	No
C00078	Trp	No	No
C00105	UMP	No	No
C00086	Urea	No	No
C00183	Val	No	No
C00041	Ala	Yes	Yes
C00956	alpha-Aminoadipate	Yes	No
C00120	Biotin	Yes	No
C00158	Citrate	Yes	Yes
C00571	Cyclohexylamine	Yes	No
C00097	Cys	Yes	No
C00122	Fumarate	Yes	Yes
C00051	Glutathione(red)	Yes	No
C00037	Gly	Yes	Yes
C00093	Glycerophosphate	Yes	No
C00262	Hypoxanthine	Yes	No
C00294	Inosine	Yes	No
C00711	Malate	Yes	No
C02704	Methyl sulfate	Yes	No
C00153	Nicotinamide	Yes	No
C02571	o-Acetylcarnitine	Yes	No
C00199	Ru5P	Yes	Yes
C00019	SAM+	Yes	No

C00315	Spermidine	Yes	No
C00042	Succinate	Yes	Yes
C00043	UDP-N-acetylglucosamine	Yes	No

Linked from mice: indicates whether the metabolite can be linked from the 24 metabolites that shared the same KEGG pathway as the differentially expressed genes in the mouse model

Table S10. KEGG pathways containing at least one differentially expressed gene and two metabolites detected in mice

KEGG pathway	Differentially expressed gene	Detected metabolite
HIF-1 signaling pathway	*Crebbp*	Lactate
		ATP
FoxO signaling pathway	*Crebbp*	ADP
		AMP
		Glu
Pentose phosphate pathway	*Pgls*	6-Phosphogluconate
		S7P
		Ru5P
		3PG
Carbon metabolism	*Pgls*	F6P
		3PG
		6-Phosphogluconate
		F1,6P
		S7P
		Ru5P
		PEP
		Gly
		Ser
		Fumarate
		Succinate
		Citrate
		Asp
		Ala
Rap1 signaling pathway	*Fgf15*	GTP
		GDP
Ras signaling pathway	*Fgf15*	GTP
		GDP
Alcoholism	*H2afx*	Tyr
		Adenosine
		Glu

Table S11. Sample Information of the rhesus macaques used in CE-MS measurements

Sample ID	Species	Tissue	Sex	Age		Experiment	PMD (hour)	Weight
				Year	Day			
F0908	Macaque	PFC	M	0	-36	CE-MS	<0.3	9.8
F0910	Macaque	PFC	M	0	-35	CE-MS	<0.3	11
F0906	Macaque	PFC	M	0	-30	CE-MS	<0.3	13.8
NB0901	Macaque	PFC	M	0	1	CE-MS	<0.3	10.9
NB0904	Macaque	PFC	F	0	2	CE-MS	<0.3	10.2
704	Macaque	PFC	M	0	20	CE-MS	<0.3	10.1
703	Macaque	PFC	M	0	22	CE-MS	<0.3	9.8
702	Macaque	PFC	M	0	23	CE-MS	<0.3	11.5
70175	Macaque	PFC	M	0	151	CE-MS	<0.3	9.8
70141	Macaque	PFC	M	0	153	CE-MS	<0.3	10.2
70133	Macaque	PFC	M	0	207	CE-MS	<0.3	10.8
6403	Macaque	PFC	M	0	215	CE-MS	<0.3	10.2
6237	Macaque	PFC	M	0	353	CE-MS	<0.3	10.7
61569	Macaque	PFC	M	1	80	CE-MS	<0.3	9.6
51087	Macaque	PFC	M	1	170	CE-MS	<0.3	12.1
51373	Macaque	PFC	M	1	242	CE-MS	<0.3	15
51469	Macaque	PFC	M	1	294	CE-MS	<0.3	9.9
51095	Macaque	PFC	M	2	9	CE-MS	<0.3	9.9
4093	Macaque	PFC	M	3	40	CE-MS	<0.3	12.6
50715	Macaque	PFC	M	3	80	CE-MS	<0.3	10
m0507910	Macaque	PFC	F	4	227	CE-MS	<0.3	10
B00051	Macaque	PFC	M	6	165	CE-MS	<0.3	11
99057	Macaque	PFC	M	8	16	CE-MS	<0.3	12.4
MM0906	Macaque	PFC	F	8	34	CE-MS	<0.3	10.1
5917	Macaque	PFC	M	10	291	CE-MS	<0.3	9.6
96007	Macaque	PFC	M	10	328	CE-MS	<0.3	10.3
m9704920	Macaque	PFC	F	12	263	CE-MS	<0.3	11.9
9605799	Macaque	PFC	M	14	297	CE-MS	<0.3	11.9
92095	Macaque	PFC	M	14	349	CE-MS	<0.3	11.9
92107	Macaque	PFC	M	15	3	CE-MS	<0.3	13.2
86023	Macaque	PFC	M	21	8	CE-MS	<0.3	13.8
9807483	Macaque	PFC	M	12	227	CE-MS	4	12.7
9807483	Macaque	PFC	M	12	227	CE-MS	4	9.6
9804105	Macaque	PFC	M	12	324	CE-MS	6	10.1

PMD: Postmortem delay. Lines highlighted with blue represent the two individuals with substantial PMD, one of them having two replicates.

Weight: sample weight (mg) for CE-MS experiment.

References

1. Sugimoto M, *et al.* (2012) Non-targeted metabolite profiling in activated macrophage secretion. *Metabolomics* 8(4):624-633.
2. Dobin A, *et al.* (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1):15-21.
3. He Z, Bammann H, Han D, Xie G, & Khaitovich P (2014) Conserved expression of lincRNA during human and macaque prefrontal cortex development and maturation. *RNA* 20(7):1103-1111.
4. Somel M, *et al.* (2009) Transcriptional neoteny in the human brain. *Proceedings of the National Academy of Sciences* 106(14):5743-5748.
5. Kim D, *et al.* (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology* 14(4):R36.
6. Yuan Y, *et al.* (2011) Development and application of a modified dynamic time warping algorithm (DTW-S) to analyses of primate brain expression time series. *BMC bioinformatics* 12(1):347.
7. Fu X, *et al.* (2011) Rapid metabolic evolution in human prefrontal cortex. *Proc Natl Acad Sci U S A* 108(15):6181-6186.
8. de Melo J, *et al.* (2008) Dlx2 homeobox gene transcriptional regulation of Trkb neurotrophin receptor expression during mouse retinal development. *Nucleic Acids Res* 36(3):872-884.
9. Colasante G, *et al.* (2008) Arx is a direct target of Dlx2 and thereby contributes to the tangential migration of GABAergic interneurons. *J Neurosci* 28(42):10674-10686.
10. Paina S, *et al.* (2011) Wnt5a is a transcriptional target of Dlx homeogenes and promotes differentiation of interneuron progenitors in vitro and in vivo. *J Neurosci* 31(7):2675-2687.
11. Iyer AK, Miller NL, Yip K, Tran BH, & Mellon PL (2010) Enhancers of GnRH transcription embedded in an upstream gene use homeodomain proteins to specify hypothalamic expression. *Mol Endocrinol* 24(10):1949-1964.
12. Le TN, *et al.* (2007) Dlx homeobox genes promote cortical interneuron migration from the basal forebrain by direct repression of the semaphorin receptor neuropilin-2. *J Biol Chem* 282(26):19071-19081.
13. Kim HD, *et al.* (2011) Class-C SOX transcription factors control GnRH gene expression via the intronic transcriptional enhancer. *Mol Endocrinol* 25(7):1184-1196.
14. Tanaka S, *et al.* (2004) Interplay of SOX and POU factors in regulation of the Nestin gene in neural primordial cells. *Mol Cell Biol* 24(20):8834-8846.
15. Wang Y, Lin L, Lai H, Parada LF, & Lei L (2013) Transcription factor Sox11 is essential for both embryonic and adult neurogenesis. *Dev Dyn* 242(6):638-653.
16. Mu L, *et al.* (2012) SoxC transcription factors are required for neuronal differentiation in adult hippocampal neurogenesis. *J Neurosci* 32(9):3067-3080.
17. Shim S, Kwan KY, Li M, Lefebvre V, & Sestan N (2012) Cis-regulatory control of corticospinal system development and evolution. *Nature* 486(7401):74-79.
18. Jankowski MP, Cornuet PK, McIlwrath S, Koerber HR, & Albers KM (2006) SRY-box containing gene 11 (Sox11) transcription factor is required for neuron survival and neurite growth. *Neuroscience* 143(2):501-514.
19. Hide T, *et al.* (2009) Sox11 prevents tumorigenesis of glioma-initiating cells by inducing neuronal differentiation. *Cancer Res* 69(20):7953-7959.

20. Bergsland M, Werme M, Malewicz M, Perlmann T, & Muhr J (2006) The establishment of neuronal properties is controlled by Sox4 and Sox11. *Genes Dev* 20(24):3475-3486.
21. Bhattaram P*, et al.* (2010) Organogenesis relies on SoxC transcription factors for the survival of neural and mesenchymal progenitors. *Nat Commun* 1:9.