Rapid metabolic evolution in human prefrontal cortex -- Supplementary Information

Supplementary Analysis on Diet and Exercise Effects

To estimate the extent to which the species-specific metabolic changes might be caused by environmental factors, we screened the literature for the studies of diet and exercise effects on metabolite levels. We identified two studies that reported significant dietary effect in humans and mice (1, 2) and five studies that reported significant exercise effects in humans and rats (3-7). In total, 37 metabolites were reported to be affected by diet and 71 by exercise. An additional 28 metabolites were found to be unaffected by exercise. Out of these, 8, 14 and 9, respectively, were detected in our dataset (Table S7). Out of the 8 diet-related metabolites, 5 overlapped with metabolites showing species-specific profiles in PFC and 4 overlapped with metabolites showing species-specific profiles in CBC (Table S8). Out of the 14 exercise-related metabolites, 7 overlapped with metabolites showing species-specific profiles in PFC and 8 in CBC, with 5 found in both brain regions. Similar extent of overlap could be found for metabolites shown not to be affected by exercise.

To test any influence of diet or exercise on metabolite differences among species, we compared metabolite sets identified in our study and the above mentioned studies. Using Fisher's exact test, we did not find significant overrepresentation of diet- or exercise-affected metabolites among metabolites with species-specific concentration profiles (p>0.2 in all tests).

Among metabolites showing human-specific changes in our study, glutamate in PFC and taurine and ascorbic acid in CBC were previously reported to be affected by exercise. Further, glutamate in PFC and taurine in CBC were previously reported to be affected by diet. Differences between species detected in our study, however, did not display patterns consistent with previously reported environmental effects. For example, glutamate has been shown to decrease in concentration with increased exercise (4) or due to high-fat diet (1). In our data, however, the main difference between human and non-human

primate glutamate concentration profiles occurs at the time of birth, with glutamate concentration being lower in human newborns than that in the chimpanzee and rhesus macaque newborns. Human newborns are not expected to be more physically active than chimpanzee and macaque babies, considering the motor development in chimpanzees and macaques occurs earlier than in humans (8-11). Further, the diet in all three species before weaning does not differ markedly with respect to fat content (12, 13). Therefore, these environmental effects cannot easily explain the difference in glutamate concentration between human babies and chimpanzee or macaque babies.

Finally, we note that a previous study testing effects of human and chimpanzee diets on gene expression profiles in mice found no detectable changes in mouse brain cortex, while significant effects were observed in liver (14). This implies that at least on the gene expression level, differences between human and chimpanzee diet appear to have limited effect on brain. Thus, the environmental effects on metabolite concentration observed in the above mentioned seven studies in human urine, muscle, mouse plasma, or rat liver might actually be less prominent or absent in the brain tissue. Taken together, these results provide no indication that excess of human-specific metabolic changes found in PFC could be explained by the influence of environmental factors.

Supplementary Materials and Methods

Samples

Human samples were obtained from the NICHD Brain and Tissue Bank for Developmental Disorders at the University of Maryland, the Netherlands Brain Bank, and the Chinese Brain Bank Center. Informed consent for use of the human tissues for research was obtained in writing from all donors or their next of kin. All subjects were defined as normal controls by forensic pathologists at the corresponding brain bank. All subjects suffered sudden death with no prolonged agonal state. Chimpanzee samples were obtained from the Yerkes Primate Center, GA, USA, the Anthropological Institute & Museum of the University of Zürich-Irchel, Switzerland, and the Biomedical Primate Research Centre, the Netherlands (eight Western chimpanzees, one Central/Eastern, and three of unknown origin). Rhesus macaque samples were obtained from the Suzhou

Experimental Animal Center, China. All non-human primates used in this study suffered sudden deaths for reasons other their participation in this study and without any relation to the tissue used.

For the metabolite analysis, approximately 100 mg of brain tissue was dissected from the frozen postmortem brain tissue on dry ice. For prefrontal cortex samples, tissue dissections were taken from the frontal part of the superior frontal gyrus, a cortical region approximately corresponding to Brodmann area 10. For cerebellar cortex samples, tissue dissections were taken from the lateral part of the cerebellar hemispheres. For all samples we took care to dissect gray matter only. The tissue samples were powdered using mortar and pestle cooled by dry ice prior to the metabolite extraction procedure. For protein analysis, an adjacent 100 mg of brain tissue were dissected from a subset of human and rhesus macaque individuals. For chimpanzees, all 11 individuals, plus an additional one, were used in the protein analysis (Table S1).

GC-MS experiment

Metabolites were extracted from the frozen tissue powder by a methanol:water:chloroform (2.5:1:1 (v/v/v)) extraction. In brief 100 mg of frozen powdered tissue material was re-suspended 1 mL extraction solution containing 0.1 µg mL^{-1} of U-$^{13}C_6$-sorbitol. The samples were incubated for 10 min at 4°C on an orbital shaker. This step was followed by ultrasonication in a bath-type sonicator for 10 min at room temperature. Finally the unsoluble tissue material was pelleted by a centrifugation step (5 min; 14,000 g) and the supernatant transferred to a fresh 2 mL Eppendorf tube. To separate the organic from the aqueous phase 300 µL H_2O and 300 µL chloroform were added to the supernatant, vortexed and centrifuged (2 min; 14,000 g). Subsequently, 200 µL of the upper, aqueous phase were collected and concentrated to complete dryness in a speed vacuum at room temperature. Extract derivatization and GC-MS measurements was performed according to (15).

The obtained metabolite concentration values (apex height of the quantitative compound identifier mass) were normalized within each sample to the abundance of an internal

standard (^{13}C sorbitol) and log2 transformed. To avoid negative values, metabolite/standard ratios were scaled up by factor 10,000 prior to log2 transformation. Next, for each species and each brain region, we filtered metabolites containing more than 20% missing values. For the remaining metabolites, we estimated missing values using k-th neighbors algorithm implemented in "knn" function in R package "EMV". The number of neighboring rows to estimate the missing values was set to five.

In addition to ^{13}C sorbitol, we used two metabolites (unknown 22 and unknown 85) with the lowest concentration variance across the full range of GC-MS profiles to independently normalize the dataset. The methods of filtering metabolites and missing-value estimation were the same as ^{13}C sorbitol normalization.

Protein expression profiling

Proteins were extracted from 100 mg of frozen brain tissue from 12 human, 12 chimpanzee, and 12 rhesus brain samples. For each species, samples were processed in two batches with similar age distribution between the batches (Table S1). The protein extraction, identification and quantification procedure was carried out as described previously (16, 17). Briefly, tissue samples were homogenized in ice-cold lysis buffer (8 M urea, 4% CHAPS, 65 mM DTT, 40 mM Tris, cocktail protease inhibitor, 100 mg of tissue/1 ml) using an electric homogenizer. The resulting protein solutions were sonicated on ice for a total of 3 minutes and then centrifugated at 25,000g for 1 hour at 4°C to remove DNA, RNA and other cell debris. Next, the protein supernatants were precipitated using 5× volumes of precipitation solution (ethanol: acetone: acetic acid = 50:50:0.1, volume ratio) at 4°C overnight, centrifugated and resolubilized in denaturing buffer [6 M guanidine hydrochloride, 100mM Tris, cocktail protease inhibitor, phosphatase inhibitors (1mM sodium orthovanadate and 1mM sodium fluoride), pH 8.3]. Protein concentration was determined using the Bradford assay. Next, 600μg of protein from each sample was reduced with DTT (100μg / 1μl 1M DTT), alkylated with IAA (100μg / 2μl 1M IAA), and precipitated again at 4°C overnight. After centrifugation, the resulting precipitates were resolubilized in digestion buffer (100mM ammonium

bicarbonate) and incubated with Trypsin (enzyme:protein = 1:40, mass ratio) at 37°C for 20 hours, followed by ultrafiltration and lyophilization.

Each peptide sample was resolubilized in 50µl SCX loading buffer, loaded on a SCX (Strong Cation Exchange) column (Column Technology Inc., CA, USA) and eluted using a pH continuous gradient buffer (from pH 2.5-8.5), resulting in 10 fractions. Each of these fractions was then automatically loaded on one of two RP (Reversed Phase) alternative trap columns, by switching to the other RP column every 3 hours using a pH continuous gradient buffer. Analysis was performed on the LTQ mass spectrometer equipped with a metal needle electrospray interface mass spectrometer (ThermoFinnigan, San Jose, CA, USA) in a data-dependent collection model (each full scan followed by ten MS/MS scans of most intense ions).

We searched the mass spectra data of each sample. The peptides were identified by searching against an IPI human database (IPI human v3.61) and a decoy database containing reversed protein sequences separately, using the SEQUEST program in BioWorks™ 3.2 software suite. A mass tolerance of 3.0 Da and one missed cleavage site of trypsin were allowed. Cysteine carboxyamidomethylation was set as static modification and methionine oxidation was set as variable modification. The output of SEQUEST was filtered using the following cutoffs: Xcorr > 1.9 for charge +1; Xcorr > 2.2 for charge +2; Xcorr > 3.75 for charge +3 and delta CN above 0.1 for all the charges. All output results were filtered and integrated to proteins by an in-house software "BuildSummary". Using the proteins from the decoy database as the estimation of the false discovery in the real database, at a cutoff of a minimum total of 20 peptide counts among all 12 samples to filter the proteins, the false discovery rate (FDR) of protein identification was estimated to be 3%. We quantified the protein expression by counting the number of identified peptides. Then, IPI peptide IDs were mapped to Ensembl gene IDs using the Ensembl database (v56). Peptides mapping to multiple genes were excluded. For Ensembl genes with multiple IPI peptide IDs, we calculated expression levels as the mean of all peptides assigned to this gene. The expression values were then log2 transformed and quantile normalized. Only proteins with total peptide counts ≥10

were used. The processed protein expression dataset is available at http://www.picb.ac.cn/Comparative/data_methods/.

Identification of metabolites affected by postmortem delay

Here, we compared the abundance of each metabolite in the CBC and PFC between the rhesus macaque that was dissected and frozen 5 hours after death and other macaques that all had ≈20 minutes postmortem interval. To incorporate metabolic changes with age, we first fitted spline curves with three degrees of freedom to each metabolite's data in a given brain region, excluding the macaque with prolonged postmortem delay. If, for both replicates, the absolute value of residuals calculated based on the fitted curve were greater for the macaque with prolonged postmortem delay than 1.96*standard deviation of the residuals of all other individuals, we classified this metabolite as postmortem-delay affected metabolite (Figure S3). For the metabolites not affected by postmortem delay, we fitted spline curves with three degrees of freedom to the concentration profiles in all individuals, including the macaque with prolonged postmortem delay. The residuals of metabolite concentration measurements were calculated by subtracting the value interpolated from the fitted curve from the actual measurement. The variance of residuals distribution in a macaque with long postmortem delay and other macaques were compared using F-test (Figure S4).

PCA analysis

For the PCA analysis, we merged the replicate measurements by taking the average (mean). The PCA analysis was performed using the "prcomp" function in the R package "stats".

Method to calculate variation explained by each factor

We used Principle Variance Component Analysis (PVCA) to calculate the variance explained by species, age, brain region, and individual (18, 19). In brief, principal

components with substantial contribution to the total variance greater than 5% were used as response variables to fit a mixed linear model with different source of variation (age, species, brain region, and individual) as random effects, using "lmer" function in the R package "lme4". The model fit via restricted maximum likelihood (REML) to get the variance component estimates. The weighted average variance was then calculated based on the eigenvalues retained from the principal component analysis. In total, the first 5 principal components with combined contribution to the total variance greater than 52% were used.

Identification of age-related metabolites/proteins

We tested the effect of age on expression levels of 92 metabolites and 1,951 proteins using polynomial regression models, as described previously (17, 20). Briefly, for each metabolite/protein, we chose the best polynomial regression with age as predictor and expression level as response, using families of polynomial regression models and the "adjusted r^2" criterion (21). The significance of the chosen regression model was estimated using the F-test. The false discovery rate was calculated by random permutation of age.

Identification of pair-wise differentially expressed metabolites/proteins

The test for differential expression between a pair of species is based on analysis of covariance, or ANCOVA (21). For each of 63 and 66 age-related metabolites in CBC and PFC, respectively and, 1,271 age-related proteins in PFC identified using the method described above, we used the age model chosen in the above-described age-test and then tested if adding species-specific parameters to the model significantly improved its fit to the expression levels in the other species. The null model (no species-specific parameters) and alternative models were compared using the F-test. In order to preserve the age-structure in the data, the permutation test was performed by dividing the age-range into 5 sections, and randomly permuting species' identifiers among samples within each section. The differential expression test was performed twice on each species pair,

using either species as a reference (for choosing the age-test model). For each metabolite/protein, if both of the two tests were significant at a defined cutoff (FDR<1%), we considered it as differentially expressed between the two species. The same methods were applied to the full dataset and a subset of 11 individuals per species.

Identification of species-specific metabolites/proteins

For the metabolite data analysis based on the full set of individuals, we used chimpanzee metabolite measurements to assigned metabolic changes for the 43 and 49 metabolites with significant differences between humans and macaques in CBC and PFC, respectively, to the evolutionary lineages: the human lineage and the lineage between rhesus macaques and the common ancestor of humans and chimpanzees. To do so, we fitted the metabolite concentration data using spline curves with three degrees of freedom and interpolated metabolite levels for 103 age points, based on the union of actual age points in the three species. The distance between a pair of species was defined as the absolute value of the difference between two vectors of 103 predicted metabolite levels. We then used a one-sided Wilcoxon test to compare human-chimpanzee difference to chimpanzee-macaque difference. If the human-chimpanzee difference was significantly greater than chimpanzee-macaque difference, we classified these metabolites as human-specific. If the chimpanzee-macaque was significantly greater than human-chimpanzee difference, we classified these metabolites as macaque-specific. The rest of age-related metabolites with significant difference between humans and macaques were classified as uncategorized.

For the two 11-individuals-per-species datasets, we tested the human-chimpanzee, macaque-chimpanzee and human-macaque differences in the pair-wise comparisons as described above. If both the human-chimpanzee and human-macaque differences were significant, but the macaque-chimpanzee difference was not, we classified these metabolites as human-specific. If both the human-chimpanzee and macaque-chimpanzee differences were significant, but the human-macaque difference was not, we classified these metabolites as chimpanzee-specific. If both the human-macaque and macaque-chimpanzee differences were significant, but the human-chimpanzee difference was not,

we classified these metabolites as macaque-specific. The rest of age-related metabolites were classified as uncategorized.

For the proteins data, we used chimpanzee protein measurements to sort the 288 human-macaque protein differences in PFC, identified by the method mentioned above into four categories. We tested the human-chimpanzee and macaque-chimpanzee differential expression with the same method we used for the human-macaque test. Those proteins whose profiles were significantly different ($p<0.005$) in humans from the profiles in both chimpanzee and macaque, but not different between chimpanzee and macaque, were classified as human-specific proteins. Similar methods were used to classify the chimpanzee-specific and macaque-specific proteins. The rest of proteins were classified as uncategorized.

Types of metabolic differences among species

The differential expressed metabolites were grouped into different categories by the following two steps: (1) The ages of human and rhesus macaque were normalized to the chimpanzee lifespan-scale by dividing the sample age by the maximum lifespan of the corresponding species (human: 120 years, rhesus macaque: 40 years) and multiplying by the maximum lifespan of chimpanzee (60 years) (22). Normalized ages were used to redo the analysis to identify differential metabolites. The metabolites that did not show significant difference after lifespan normalization were classified as different due to the lifespan difference. (2) The data after removing the mean, by standardizing metabolite concentration level in each species to mean = 0 and standard deviation = 1, was used to identify differential metabolites. The metabolites that did not show significant difference after mean and variance normalization were classified as different due to the mean concentration differences. The remaining metabolites were classified as showing different patterns of metabolic change with age among species.

Heatmap and hierarchical clustering

For each of the 24 and 6 human-specific metabolites identified as described above in PFC and CBC, respectively, we interpolated metabolite concentration values at 50 time points uniformly distributed over the species' lifespan, based on the spline curves with three degrees of freedom, fitted to the metabolite measurements in each species and each brain region. All species' curves were extrapolated to the minimum and maximum age of the dataset (0 and 98 years). In a given brain region, the fitted 150 concentration values for each metabolite were normalized to mean equal zero and standard deviation equal one before clustering. The heatmap and hierarchical clustering dendrogram were drawn by the "heatmap.2" function in the R package "gplots". The clustering was based on 1-Pearson correlation coefficient, as the distance measure. For the human-specific metabolites identified in PFC, the clustering was done based on the PFC metabolite concentrations. For the human-specific metabolites identified in CBC, the hierarchical clustering was done based on the CBC metabolite concentrations.

Test relationship between metabolites and proteins

The proteins connected by a single edge to the detected metabolites in the pathways were manually collected from the KEGG database (http://www.kegg.com). We found 38 such proteins for 8 human-specific metabolites and 106 proteins for 23 non-human-specific metabolites. The 11 proteins that were shared among human-specific metabolites and non-human-specific metabolites were assigned to human specific metabolites. The Pearson correlation between profiles of metabolite-protein pairs were calculated using the 50 points interpolated from spline curves fitted to the metabolite and protein data, with three degrees of freedom. Three Pearson correlation coefficient cutoffs ($|r|>0.3$, $|r|>0.5$, $|r|>0.7$) were applied to define correlated metabolite-protein pairs. The excess of observed correlated metabolite-protein pairs was tested by random substitutions of proteins' expression profiles by expression profiles of other detected proteins 1,000 times. The *p*-value was calculated as the frequency of the event when the number of random pairs identified using a given correlation coefficient cutoff was greater or equal to the number of pairs found in the actual data. The protein expression divergence between human and chimpanzee was calculated as the Euclidian distance between the

two vectors of 50 points interpolated from spline curves fitted to the protein expression data with three degrees of freedom. The human-specific protein expression divergence was calculated as the ratio between human-macaque expression divergence and chimpanzee-macaque expression divergence. We used a one-sided Wilcoxon test to compare both the human-chimpanzee divergence measurements and human-specific divergence measurements between proteins associated with the human-specific metabolites and proteins associated with non-human-specific metabolites.

KEGG pathway analysis

The xml files of human KEGG pathways were downloaded from ftp://ftp.genome.jp /pub/kegg/xml/kgml and parsed by R package "KEGGgraph". The human-specific metabolites and detected proteins were mapped to the KEGG pathways. We only considered the pathways containing at least one human-specific metabolite and more than 10 detected proteins. Using these criteria, we obtained 13 such KEGG pathways including 5 human-specific metabolites and 209 detected proteins. Then, using the proteins detected in each pathway, we calculated the human-specific protein expression divergence as described above. We then estimated the significance of these estimates by randomly assigning the same number of detected proteins to each pathway and calculating the human-specific protein expression divergence 1,000 times. The *p*-value for each pathway was calculated as the frequency of the event when the median value of the human-specific divergence in a random sub-sampling was greater or equal to the one found in the real pathway. To estimate FDR when the particular test is called significant, we applied the Storey and Tibshirani approach to calculate the q-value among the *p*-values from the 13 KEGG pathways, using R package "qvalue" (23). This resulted in three human-specific metabolites associated with four pathways showing significant excess of the human-specific protein expression divergence. The Pearson correlations between human-specific metabolites and proteins in these four pathways were calculated using the 50 points interpolated from spline curves, fitted to the metabolite and protein data, with three degrees of freedom. The Pearson correlation coefficient cutoff was set to $|r|>0.7$.

Supplementary Figures

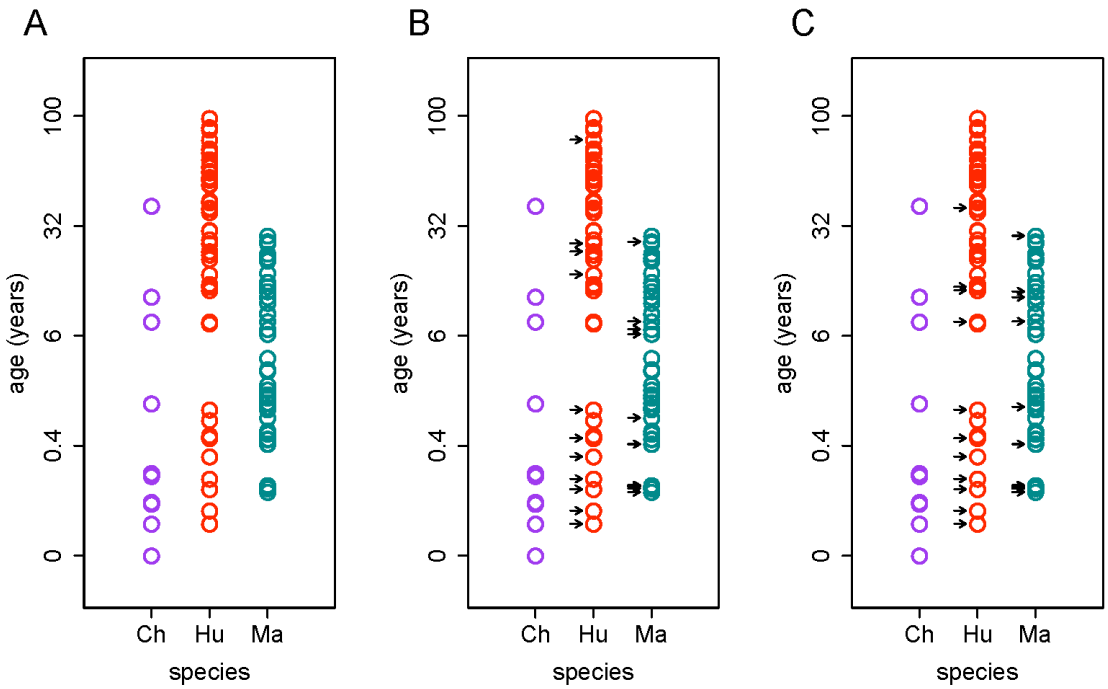


Figure S1. The sample age distributions in the three species.

A. The actual age distribution of sampled individuals in the three species (Ch-chimpanzee, Hu-human, Ma-rhesus macaque). The circles represent individuals (purple-chimpanzee, red-human, green-rhesus macaque). The y-axis shows individuals' age on the scale: (age in days)^0.25. We used this scale because it provides a nearly uniform distribution of sample ages across the lifespan. Using non-transformed age or logarithm-transformed age leads to diminished resolution at either young or old age. B. Subset of 11 human and 11 rhesus macaque individuals sampled using the stage-of-life matching strategy (indicated by arrows): scaling ages linearly within each species to the same maximal lifespan, using 120 years for humans, 60 for chimpanzees and 40 for rhesus macaques as a reference (22). In this approach, 20 years old macaque would match 60 years old human. C. Subset of 11 human and 11 rhesus macaque individuals sampled using the chronological matching strategy (indicated by arrows): using calendar age directly, *i.e.* 20 years old macaque matching 20 years old human.

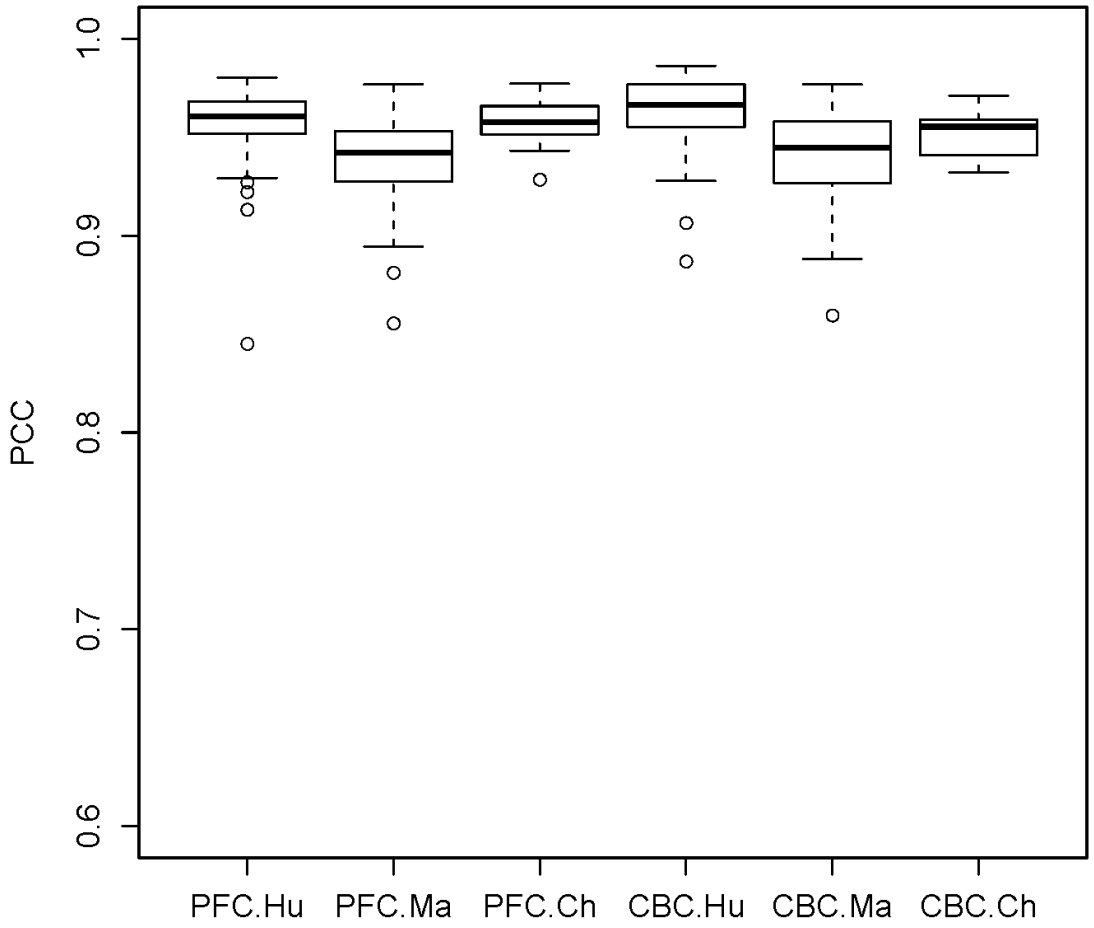
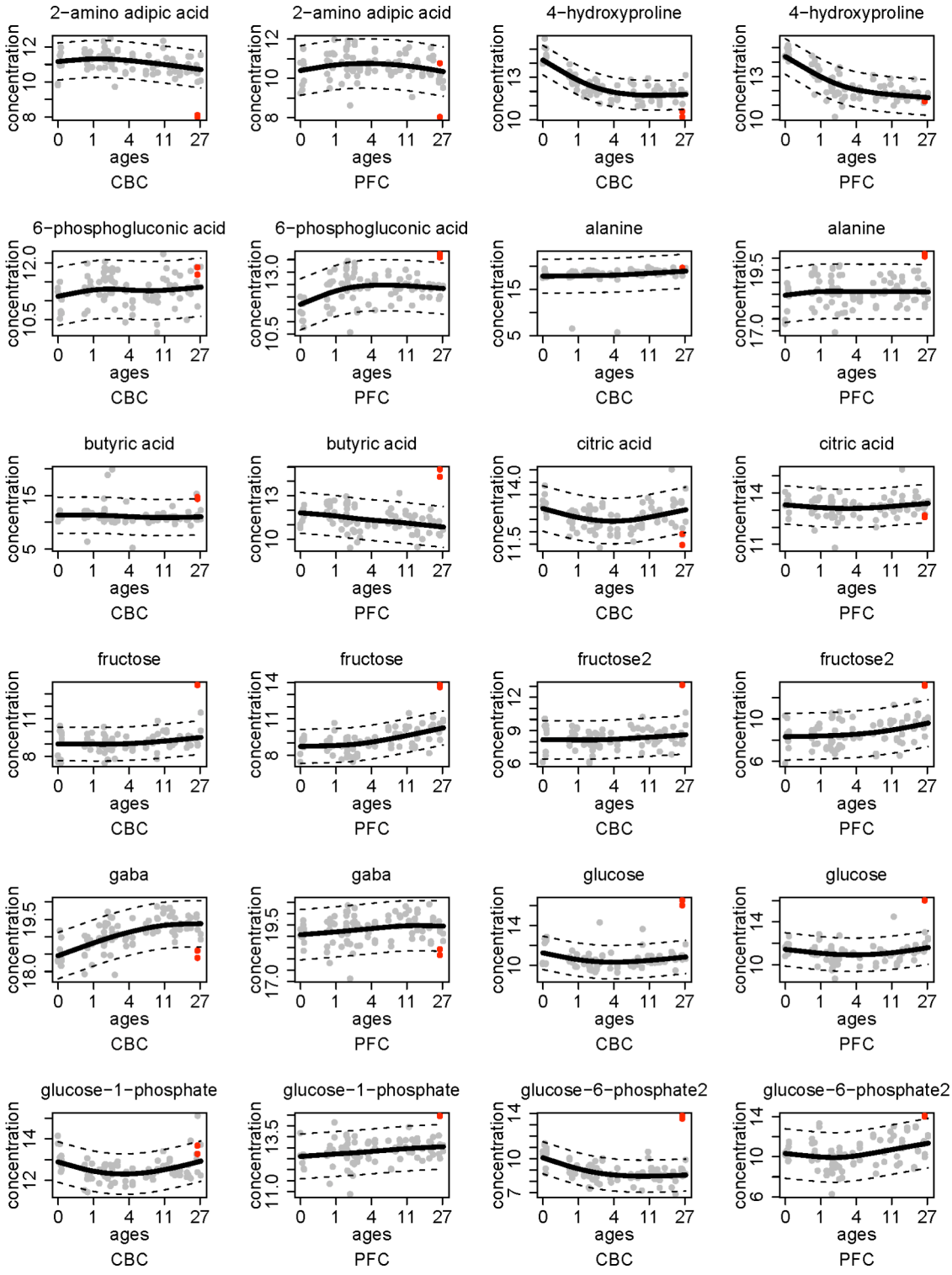
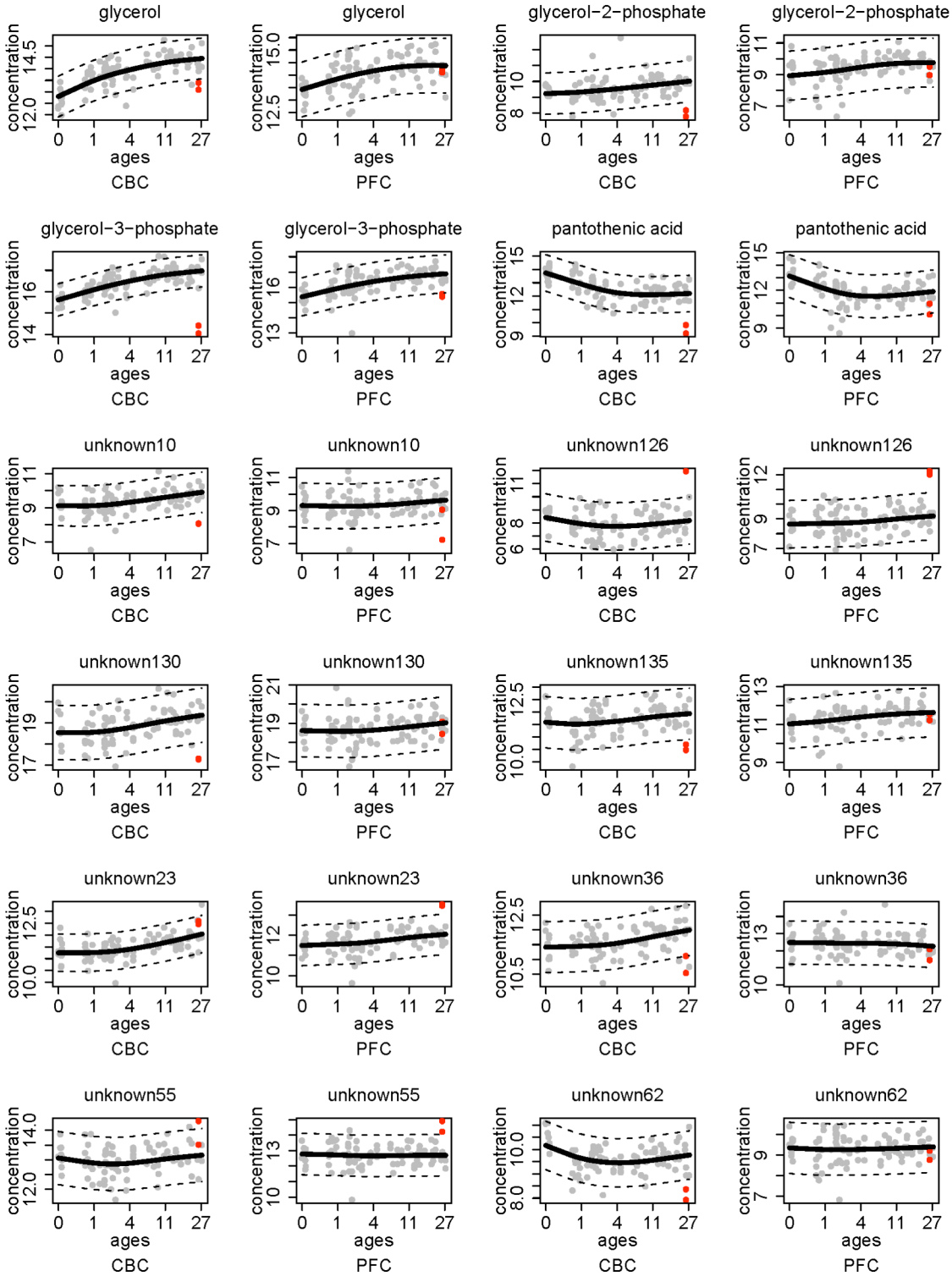


Figure S2. Correlation between technical replicates.

Sown are distributions of Pearson correlation coefficients (PCC) based on metabolite concentration measurements between technical replicates in prefrontal cortex (PFC) and cerebellar cortex (CBC) of humans (Hu), rhesus macaques (Ma) and chimpanzees (Ch). The bottom and top of each box are the 25th and 75th percentiles of the PCC distribution, and the line within the box is the median. The whiskers show 1.5 interquartile range of the lower and the upper quartiles. The data points not included between the whiskers are plotted as circles.


2−amino adipic acid
concentration
ages
CBC
2−amino adipic acid
concentration
ages
PFC
4−hydroxyproline
concentration
ages
CBC
4−hydroxyproline
concentration
ages
PFC
6−phosphogluconic acid
concentration
ages
CBC
6−phosphogluconic acid
concentration
ages
PFC
alanine
concentration
ages
CBC
alanine
concentration
ages
PFC
butyric acid
concentration
ages
CBC
butyric acid
concentration
ages
PFC
citric acid
concentration
ages
CBC
citric acid
concentration
ages
PFC
fructose
concentration
ages
CBC
fructose
concentration
ages
PFC
fructose2
concentration
ages
CBC
fructose2
concentration
ages
PFC
gaba
concentration
ages
CBC
gaba
concentration
ages
PFC
glucose
concentration
ages
CBC
glucose
concentration
ages
PFC
glucose−1−phosphate
concentration
ages
CBC
glucose−1−phosphate
concentration
ages
PFC
glucose−6−phosphate2
concentration
ages
CBC
glucose−6−phosphate2
concentration
ages
PFC


glycerol
concentration
ages
CBC
glycerol
concentration
ages
PFC
glycerol−2−phosphate
concentration
ages
CBC
glycerol−2−phosphate
concentration
ages
PFC
glycerol−3−phosphate
concentration
ages
CBC
glycerol−3−phosphate
concentration
ages
PFC
pantothenic acid
concentration
ages
CBC
pantothenic acid
concentration
ages
PFC
unknown10
concentration
ages
CBC
unknown10
concentration
ages
PFC
unknown126
concentration
ages
CBC
unknown126
concentration
ages
PFC
unknown130
concentration
ages
CBC
unknown130
concentration
ages
PFC
unknown135
concentration
ages
CBC
unknown135
concentration
ages
PFC
unknown23
concentration
ages
CBC
unknown23
concentration
ages
PFC
unknown36
concentration
ages
CBC
unknown36
concentration
ages
PFC
unknown55
concentration
ages
CBC
unknown55
concentration
ages
PFC
unknown62
concentration
ages
CBC
unknown62
concentration
ages
PFC

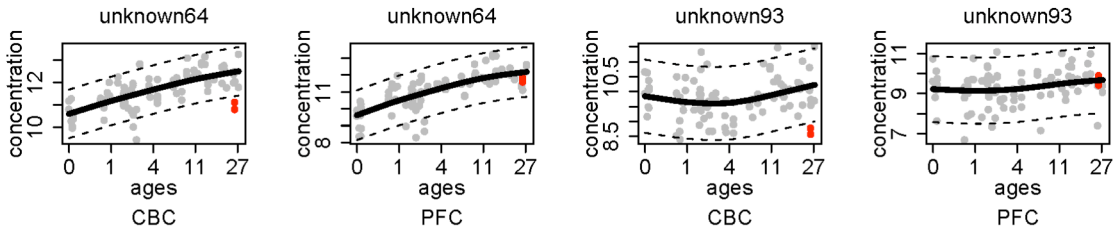


Figure S3. Concentration profiles of metabolites affected by postmortem delay.

The points represent metabolite concentrations in CBC (odd panels) and PFC (even panels) of macaque individuals with short (<20 minutes) postmortem delay (gray) and one macaque individual with long (≈5 hours) postmortem delay (red). The replicate measurements for each individual are shown separately. The lines are the spline curves fitted to the data points excluding the long postmortem delay individual with three degrees of freedom. The x-axis shows individual's ages on the scale: (age in days)^0.25. The y-axis shows the normalized GC-MS measurements representing the metabolite concentrations. The titles show metabolite annotation. "Unknown" stands for unannotated metabolites. The brain region identity is displayed below the x-axis labels.

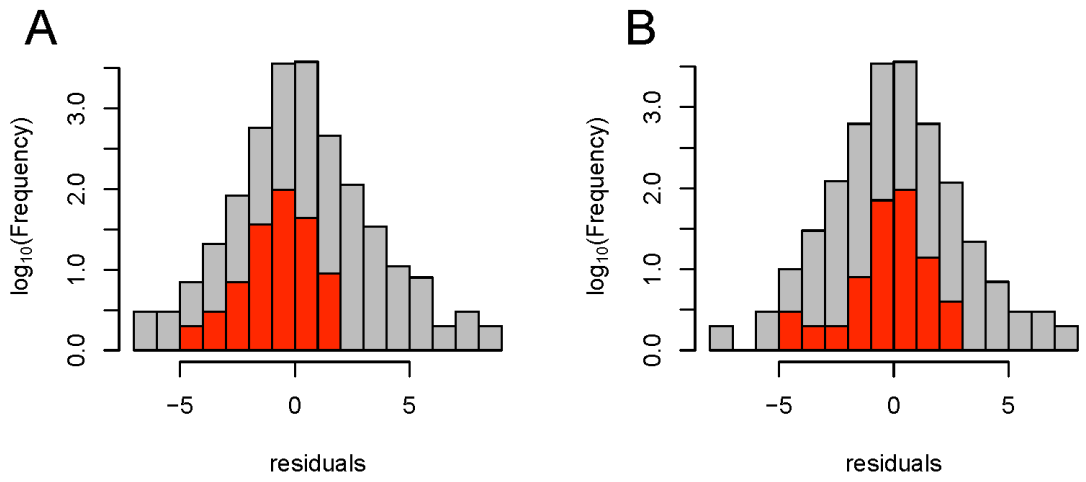


Figure S4. Variation among metabolites not affected by postmortem delay.

Shown are the distributions of residuals of metabolite concentrations measured for metabolites not affected by postmortem delay in a macaque with long postmortem delay (red) and other macaques (gray) in CBC (A) and in PFC (B). Residuals were calculated residuals to the spline curve fitted to the full metabolite data series with three degrees of freedom. The variance of the red and gray distributions was compared by the F-test. There were no significant differences in residuals distributions: the *p*-value equaled 0.254 in CBC and 0.567 in PFC. Thus, metabolites not affected by postmortem delay according to our criteria (see Materials and Methods) did not display greater variation in a macaque with 5 hours postmortem delay than in other macaques with postmortem delay less than 20 minutes (see Materials and Methods)

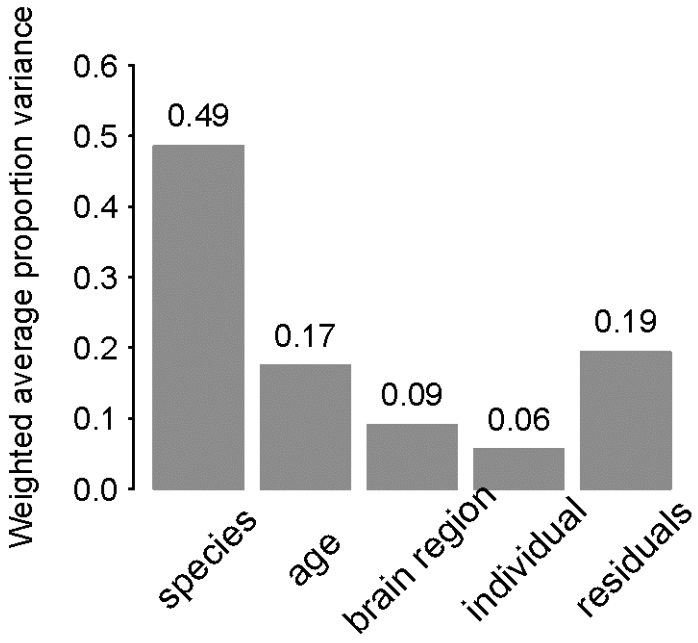


Figure S5. The proportion of variance explained by species, age, brain region, and individual effects in metabolite data.

The normalized metabolite data with 92 detected metabolites from the three species were used in the principal component analysis. Principal components with substantial contribution to the total variance greater than 5% were used as response variables to fit a mixed linear model with different source of variability (age, species, brain region and individual) as random effects (18, 19). The model was fitted via restricted maximum likelihood (REML) and was used to obtain the variance component estimates. The weighted average variance was then calculated based on the eigenvalues retained from the principal component analysis. In total, the first 5 principal components with combined contribution to the total variance >52% were used.

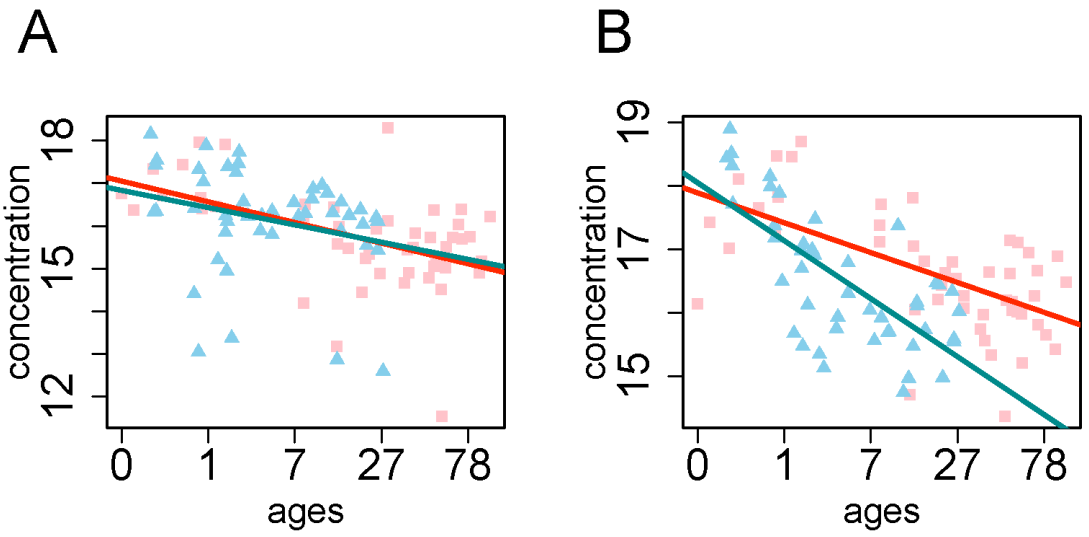
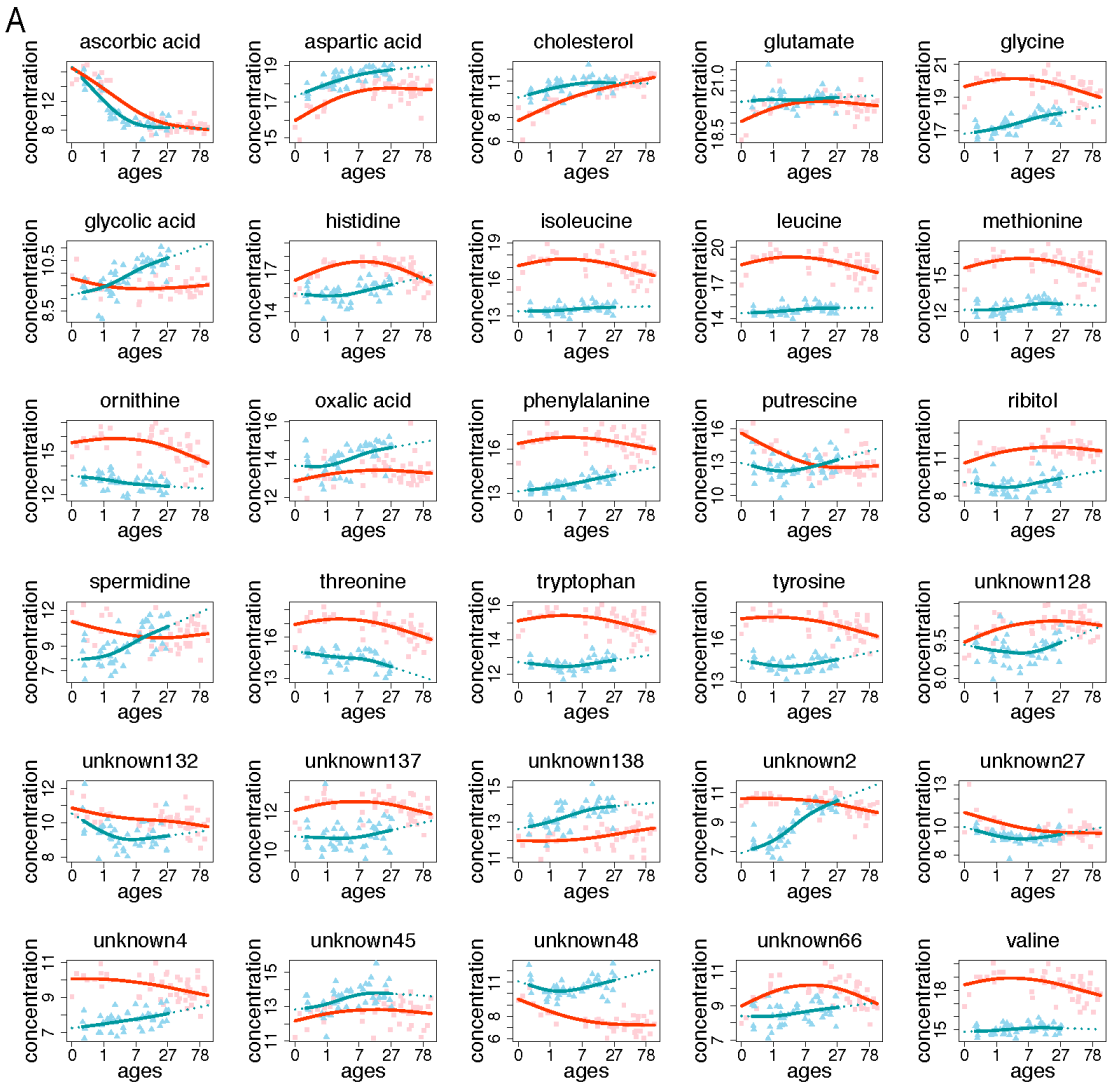


Figure S6. Concentration profiles of taurine.

The concentration profiles of taurine in PFC (A) and CBC (B) in humans (red) and rhesus macaques (blue). The points represent metabolite concentration in each individual. The lines represent fitted linear models. The x-axis shows individual's ages on the scale: (age in days)^0.25. The y-axis shows the normalized GC-MS measurements representing metabolite concentrations. The concentration of taurine shows significant decrease with age (human PFC: r=-0.37, p<0.01; human CBC: r=-0.53, p<10^{-5}; macaque PFC: r=-0.28, p<0.05; macaque CBC: r=-0.51, p<0.001).


A
ascorbic acid
aspartic acid
cholesterol
glutamate
glycine
glycolic acid
histidine
isoleucine
leucine
methionine
ornithine
oxalic acid
phenylalanine
putrescine
ribitol
spermidine
threonine
tryptophan
tyrosine
unknown128
unknown132
unknown137
unknown138
unknown2
unknown27
unknown4
unknown45
unknown48
unknown66
valine
concentration
ages
0 1 7 27 78

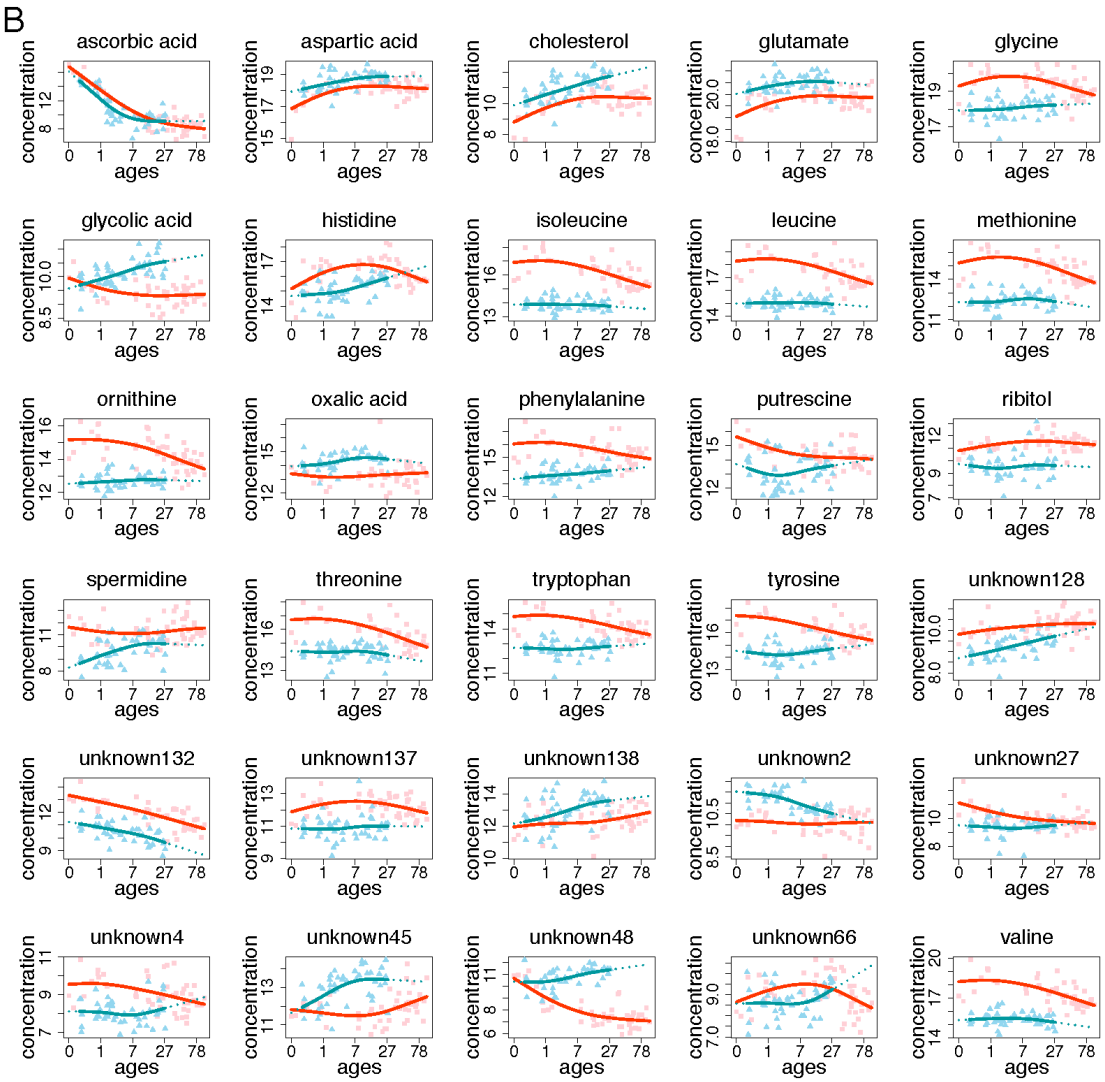


Figure S7. Concentration profiles of metabolites that differ significantly between humans and rhesus macaques in both CBC and PFC.

The profiles of 30 metabolites in CBC (A) and PFC (B) classified as significantly different between humans (red) and rhesus macaques (blue) in both brain regions. The points represent metabolite concentration in each individual. The lines are the spline curves fitted to the data points with three degrees of freedom. The x-axis shows individual's ages on the scale: (age in days)^0.25. The y-axis shows the normalized GC-MS measurements representing metabolite concentrations. The titles show metabolite annotation. "Unknown" stands for unannotated metabolites.

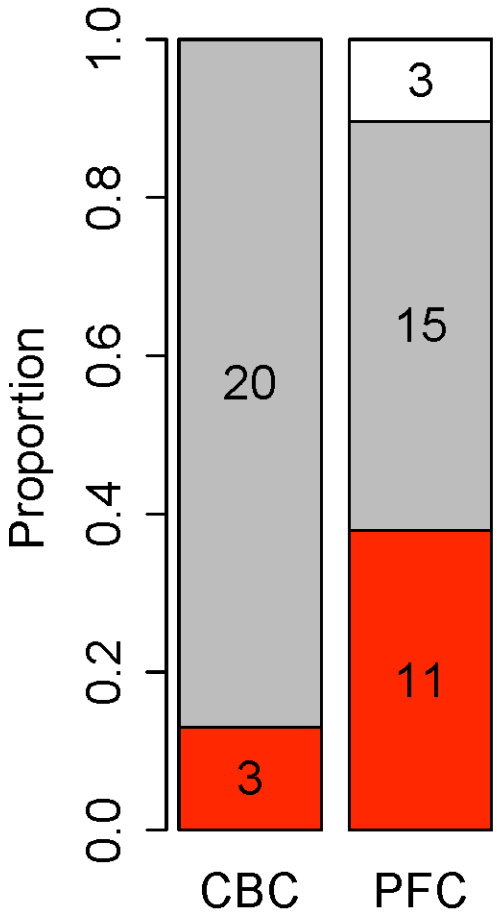


Figure S8. Proportions of annotated metabolites in the three phylogenetic categories.

The colors show the three phylogenetic categories: human-specific (red), macaque-specific (gray) and uncategorized metabolites (white). Shown are the 23 and 29 annotated metabolites with significant difference in concentration profiles between humans and rhesus macaques in CBC and PFC, respectively, identified using the full set of individuals. The numbers show numbers of metabolites in the corresponding category.

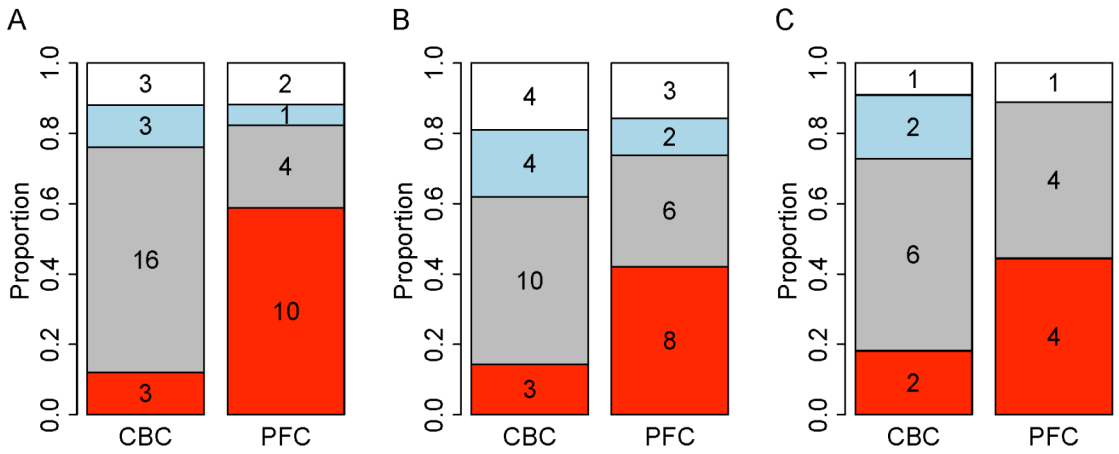


Figure S9. Proportions of metabolites in the four phylogenetic categories calculated using 11 chimpanzees and subsets of 11 human and 11 rhesus macaque individuals.

A. The distribution of metabolites in the four phylogenetic categories calculated using the subset of 11 stage-of-life matched individuals per species. B. The distribution of metabolites in the four phylogenetic categories calculated using the subset of 11 chronological matched individuals per species. C. The overlap of between the stage-of-life matching strategy and chronological matching strategy. The colors represent the four phylogenetic categories: human-specific (red), macaque-specific (gray), chimpanzee-specific (sky-blue) and uncategorized metabolites (white). The numbers indicate numbers of metabolites in the corresponding category.

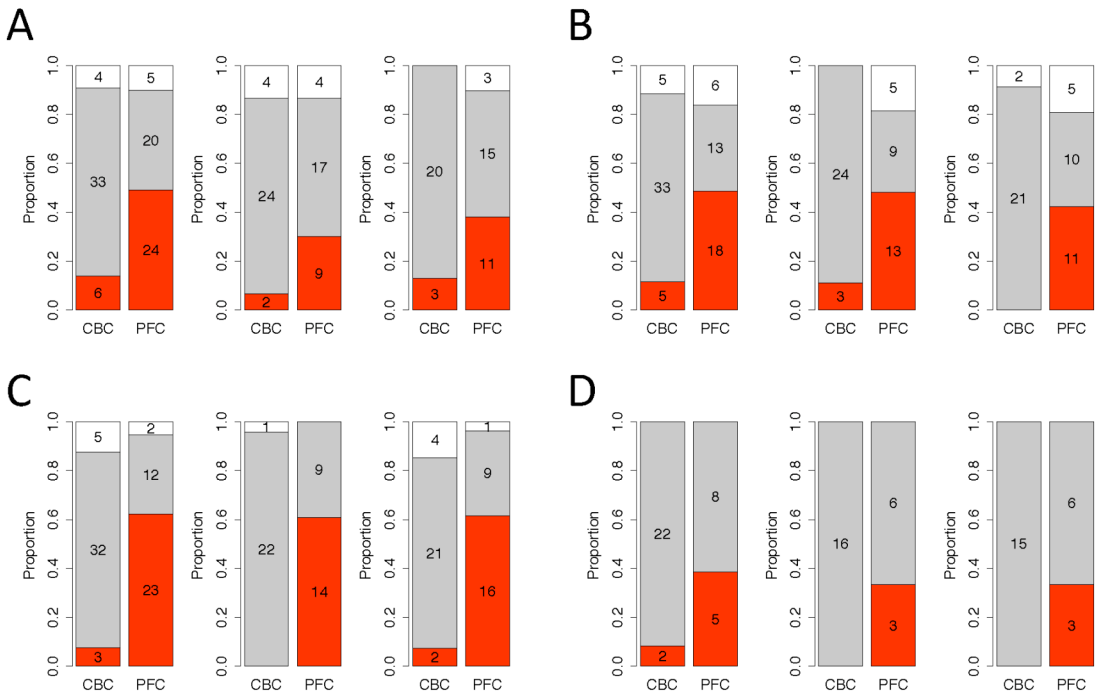


Figure S10. Proportions of metabolites in the three phylogenetic categories calculated using different normalization strategies.

A. The analysis results based on ^{13}C sorbitol normalization. B. The analysis results based on "unknown 22" normalization. C. The analysis results based on "unknown 85" normalization. D. Overlap among the results from the three normalization methods.

On each panel, shown are: (left) the proportions of metabolites with significant difference in concentration profiles between humans and macaques identified in each brain region; (middle) the proportions of metabolites with significant difference in concentration profiles between humans and macaques identified in both brain regions; (right) the proportions of annotated metabolites with significant difference in concentration profiles between humans and macaques identified in each brain region.

The colors represent the three phylogenetic categories: human-specific (red), macaque-specific (gray) and uncategorized metabolites (white). The numbers show numbers of metabolites in the corresponding category. The calculations were based on the full set of individuals.

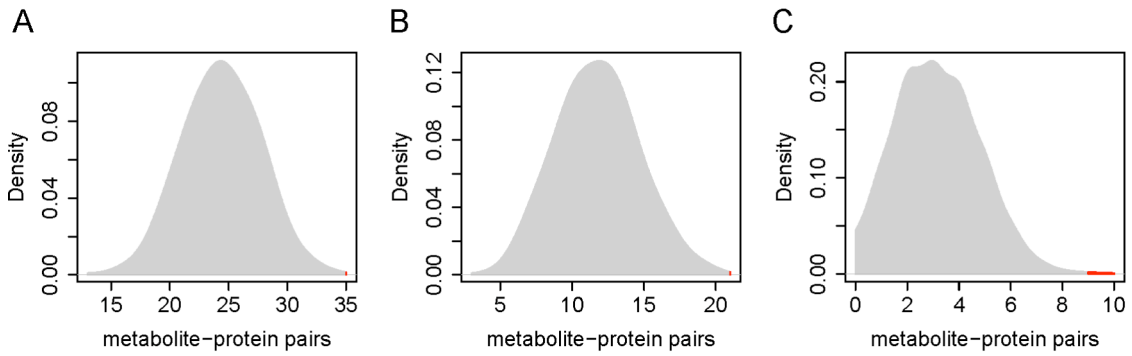


Figure S11. The numbers of metabolite-protein pairs with correlated expression profiles.

Shown are the distributions of the numbers of metabolite-protein pairs that pass a given correlation cutoff for the randomly chosen pairs (gray) and the actual number of correlated metabolite-protein pairs containing metabolites with human-specific concentration profiles (red). The Pearson correlation coefficient was calculated for 47 metabolite-enzyme pairs containing metabolites with human-specific concentration profiles and for the same number of random metabolite-protein pairs 1,000 times. The *p*-value was calculated as the frequency that the number of correlated random pairs was greater or equal to that of true pairs correlation coefficient cutoffs ($|r|$) equal 0.3 (A), 0.5 (B), and 0.7 (C). The numbers of real metabolite-protein pairs passing these cutoffs were 35, 21, and 9, respectively. The corresponding *p*-values were 0.001, 0.001, and 0.003.

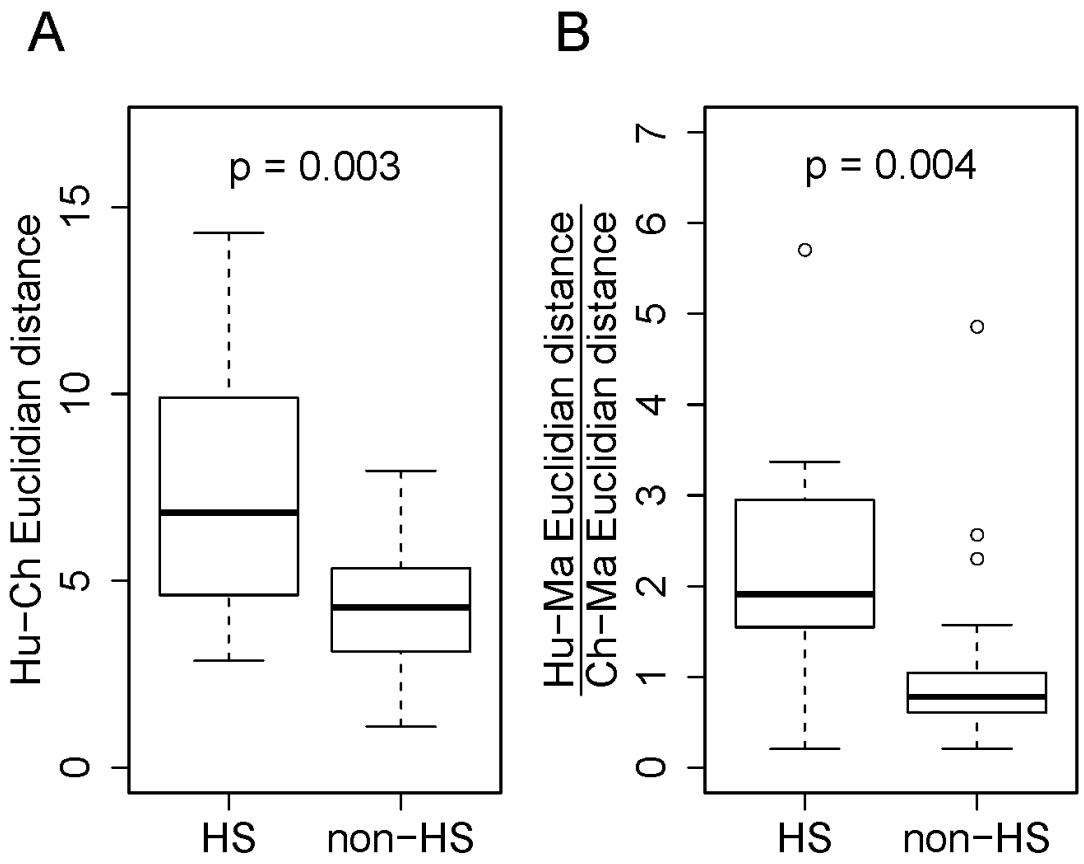


Figure S12. Expression divergence of proteins associated with human-specific and non-human-specific metabolites.

The protein's association with metabolites is based on KEGG annotation. The proteins are divided into two categories: 13 proteins associated with human-specific metabolites by a single network edge (HS) and 24 proteins associated with non-human-specific metabolites only (non-HS). A. The expression divergence between humans and chimpanzees. The divergence was calculated as Euclidian distance between protein expression profiles in humans and chimpanzees. B. The ratio of human-macaque and chimpanzee-macaque expression divergence. The expression divergence was calculated as Euclidian distance between protein expression profiles in the two species.

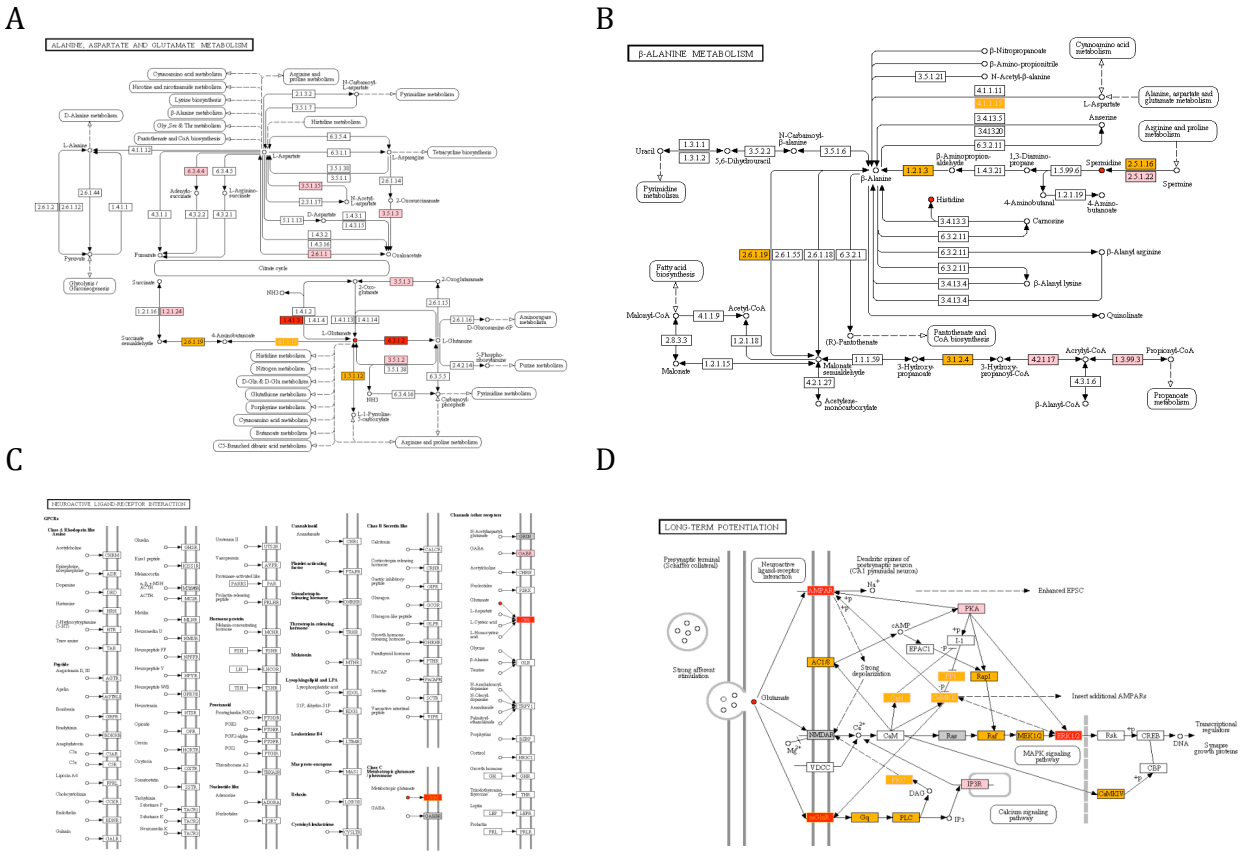


Figure S13. The four KEGG pathways with significantly greater human-specific protein expression divergence associated with three metabolites with human-specific concentration profiles.

The small red circles represent metabolites with human-specific concentration profiles: glutamate, histidine and spermidine. The boxes in the figures represent the genes in the pathways. The colors of the boxes indicate: proteins with significant human-specific expression (red), proteins with significant human-chimpanzee expression divergence (orange), proteins with significant correlation to the corresponding metabolite profile (pink), detected proteins (gray). Two type of colors (fill and border/text of boxes) are used when two type of proteins are included in the nodes. The pathways are drawn by the KEGG Mapper (http://www.genome.jp/kegg/tool/color_pathway.html).

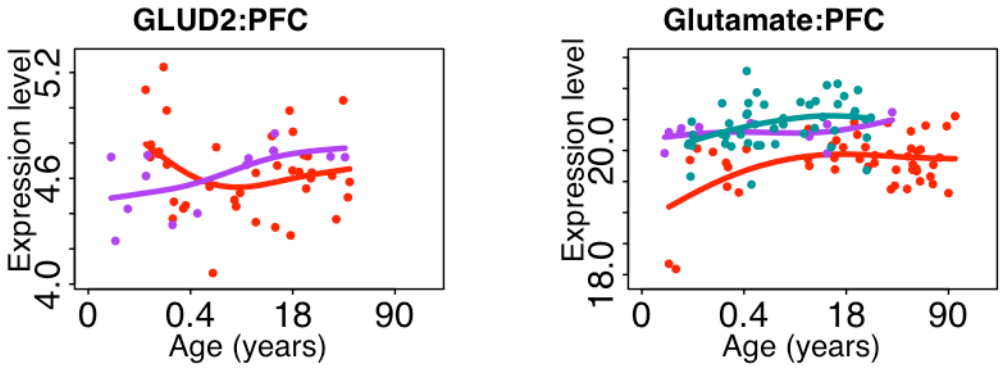


Figure S14. Expression profiles of GLUD2 mRNA and glutamate in the primate prefrontal cortex.

Colors represent species (human-red; chimpanzee-purple; macaque-blue). The points represent the expression levels in each sample. The lines are the spline curves fitted to the data points. The x-axis shows sample's ages on scale: (age in days)^(1/4).

Supplementary Tables

Table S1. Sample information.

Sample ID	Species	Gender	Age (day)	Postmortem delay (hour)	Experiment*	Batch
C1	Chimpanzee	male	0	NA	M; P	2
C2	Chimpanzee	female	1	NA	M; P	2
C3	Chimpanzee	male	7	NA	M; P	1
C4	Chimpanzee	male	8	NA	M; P	1
C5	Chimpanzee	male	39	NA	M; P	2
C6	Chimpanzee	female	45	NA	M; P	1
C7	Chimpanzee	female	522	NA	M; P	2
C8	Chimpanzee	female	2920	NA	M; P	2
C9	Chimpanzee	male	4380	NA	M; P	1
C10	Chimpanzee	male	4380	NA	M; P	2
C11	Chimpanzee	male	14600	NA	M; P	1
C12	Chimpanzee	female	16060	NA	P	1
M1	Macaque	male	16	0	M; P	1
M2	Macaque	male	20	0	M; P	2
M3	Macaque	male	22	0	M	
M4	Macaque	male	23	0	M	
M5	Macaque	male	24	0	M	
M6	Macaque	male	151	0	M	
M7	Macaque	male	153	0	M; P	1
M8	Macaque	male	179	0	M	
M9	Macaque	male	182	0	M	
M10	Macaque	male	207	0	P	2
M11	Macaque	male	215	0	M	
M12	Macaque	male	237	0	M	
M14	Macaque	male	310	0	P	2
M15	Macaque	male	353	0	M	
M16	Macaque	male	445	0	M	
M17	Macaque	male	449	0	M	
M18	Macaque	male	462	0	M	
M19	Macaque	male	471	0	M	
M20	Macaque	male	479	0	M	
M21	Macaque	male	535	0	M	
M22	Macaque	male	607	0	M	
M23	Macaque	male	647	0	M	
M24	Macaque	male	659	0	M	
M25	Macaque	male	739	0	M; P	2
M26	Macaque	male	831	0	M	
M27	Macaque	male	1135	0	M	
M28	Macaque	male	1175	0	M	
M30	Macaque	male	1487	0	M; P	1
M31	Macaque	male	2355	0	M	

M32	Macaque	male	2570	0	M	
M33	Macaque	male	2936	0	M	
M34	Macaque	male	3322	0	M	
M35	Macaque	male	3389	0	M; P	2
M36	Macaque	male	3978	0	M	
M37	Macaque	male	4361	0	M	
M38	Macaque	male	4762	0	M	
M39	Macaque	male	5131	0	M	
M40	Macaque	male	5459	0	M	
M41	Macaque	male	5478	0	M	
M42	Macaque	male	6227	0	M	
M43	Macaque	male	7391	0	M; P	1
M44	Macaque	male	7673	0	M	
M45	Macaque	male	8104	0	M; P	2
M46	Macaque	female	9291	0	M	
M47	Macaque	male	9518	0	M; P	1
M48	Macaque	female	9490	5	M	
M49	Macaque	female	10220	0	M; P	2
H1	Human	male	1	NA	M; P	1
H2	Human	male	4	NA	M; P	2
H3	Human	female	19	NA	M	
H4	Human	male	34	NA	M; P	1
H5	Human	male	94	NA	M	
H6	Human	female	186	NA	M	
H7	Human	female	196	NA	M	
H8	Human	male	204	NA	M; P	2
H9	Human	female	328	NA	M	
H10	Human	male	443	19	M	
H11	Human	male	2827	12	M	
H12	Human	male	2861	18	M	
H13	Human	male	2922	5	M; P	1
H14	Human	male	4844	5	M	
H15	Human	male	5105	13	M; P	2
H16	Human	male	5170	16	M	
H17	Human	male	5308	16	M	
H18	Human	male	6111	15	M	
H19	Human	male	7555	12	M	
H20	Human	male	8006	13	M	
H21	Human	male	8364	4	M	
H22	Human	male	9277	19	M; P	1
H23	Human	male	9743	18	M	
H24	Human	male	10885	12	M	
H25	Human	male	10948	19	M	
H26	Human	male	13587	12	M	
H27	Human	male	14101	5	M	
H29	Human	male	14309	12	M	

H30	Human	male	15206	15	M	
H31	Human	male	15519	19	M	
H32	Human	male	18406	8	M	
H33	Human	male	18615	NA	M	
H34	Human	male	19457	17	M; P	2
H35	Human	male	19566	19	M	
H36	Human	male	19930	17	M	
H37	Human	male	20030	19	M	
H38	Human	male	21050	16	M	
H39	Human	male	21204	9	M	
H40	Human	male	21365	17	M	
H41	Human	male	22452	6	M	
H42	Human	female	22265	NA	M	
H43	Human	male	24090	NA	M; P	1
H44	Human	male	25786	28	M	
H45	Human	male	26829	21	M	
H46	Human	female	26645	NA	M	
H47	Human	male	29200	NA	M; P	2
H48	Human	male	32120	NA	M; P	1
H49	Human	female	32850	NA	M	
H50	Human	male	35770	NA	M; P	2

* M: metabolite --- GC-MS experiment; P: proteomics --- label free mass-spectrometry

Table S2. List of detected metabolites.

The 1-labels indicate metabolites detected in at least 80% of samples in a corresponding tissue in a given species. The 0-labels indicate metabolites not passing the detection cutoff.

	CBC			PFC		
	chimpanzee	human	macaque	chimpanzee	human	macaque
3-hydroxybutyric acid	1	1	1	1	1	1
glycerol	1	1	1	1	1	1
leucine	1	1	1	1	1	1
isoleucine	1	1	1	1	1	1
glycine	1	1	1	1	1	1
benzoic acid	1	1	1	1	1	1
serine	1	1	1	1	1	1
unknown16	1	1	1	1	1	1
succinic acid	1	1	1	1	1	1
threonine	1	1	1	1	1	1
unknown138	1	1	1	1	1	1
fumaric acid	1	1	1	1	1	1
unknown20	1	1	1	1	1	1
unknown23	1	1	1	1	1	1
erythronic acid	1	1	1	1	1	1
4-hydroxyproline	1	1	1	1	1	1
gaba	1	1	1	1	1	1
threonic acid	1	1	1	1	1	1
methionine	1	1	1	1	1	1
creatinine	1	1	1	1	1	1
oxoproline	1	1	1	1	1	1
glutamate	1	1	1	1	1	1
nicotinamide	1	1	1	1	1	1
putrescine	1	1	1	1	1	1
dodecanoic acid	1	1	1	1	1	1
glutamine	1	1	1	1	1	1
phenylalanine	1	1	1	1	1	1
unknown38	1	1	1	1	1	1
unknown45	1	1	1	1	1	1
unknown48	1	1	1	1	1	1
glucose-1-phosphate	1	1	1	1	1	1
ornithine	1	1	1	1	1	1
aspartic acid	1	1	1	1	1	1
glycerol-3-phosphate	1	1	1	1	1	1
glucose	1	1	1	1	1	1
citric acid	1	1	1	1	1	1
unknown55	1	1	1	1	1	1
dehydroascorbic acid	1	1	1	1	1	1
unknown62	1	1	1	1	1	1
histidine	1	1	1	1	1	1
pantothenic acid	1	1	1	1	1	1
unknown71	1	1	1	1	1	1

hexadecanoic acid	1	1	1	1	1	1
unknown135	1	1	1	1	1	1
unknown128	1	1	1	1	1	1
octadecenoic acid	1	1	1	1	1	1
6-phosphogluconic acid	1	1	1	1	1	1
tryptophan	1	1	1	1	1	1
uridine	1	1	1	1	1	1
alanine	1	1	1	1	1	1
butyric acid	1	1	1	1	1	1
unknown40	1	1	1	1	1	1
tetradecanoic acid	1	1	1	1	1	1
unknown91	1	1	1	1	1	1
cholesterol	1	1	1	1	1	1
hydroxypyridine	1	1	1	1	1	1
unknown5	1	1	1	1	1	1
unknown7	1	1	1	1	1	1
3-hydroxypyridine	1	1	1	1	1	1
ribitol	1	1	1	1	1	1
fructose	1	1	1	1	1	1
unknown66	1	1	1	1	1	1
unknown137	1	1	1	1	1	1
unknown36	1	1	1	1	1	1
unknown8	1	1	1	1	1	1
glycerol-2-phosphate	1	1	1	1	1	1
tyrosine	1	1	1	1	1	1
unknown79	1	1	1	1	1	1
pyruvic acid	1	1	1	1	1	1
unknown130	1	1	1	1	1	1
unknown27	1	1	1	1	1	1
unknown136	1	1	1	1	1	1
heptadecanoic acid	1	1	1	1	1	1
unknown1	1	1	1	1	1	1
unknown19	1	1	1	1	1	1
valine	1	1	1	1	1	1
unknown108	1	1	1	1	1	1
oxalic acid	1	1	1	1	1	1
decanoic acid	1	1	1	1	1	1
unknown93	1	1	1	1	1	1
unknown132	1	1	1	1	1	1
unknown3	1	1	1	1	1	1
taurine	1	1	1	1	1	1
unknown64	1	1	1	1	1	1
tetracosanoic acid	1	1	1	1	1	1
glycolic acid	1	1	1	1	1	1
unknown126	1	1	1	1	1	1
unknown85	1	1	1	1	1	1
unknown6	1	1	1	1	1	1
unknown33	1	1	1	1	1	1
unknown49	1	1	1	1	1	1
unknown2	1	1	1	1	1	1
unknown22	1	1	1	1	1	1

unknown30	1	1	1	1	1	1
unknown4	1	1	1	1	1	0
unknown35	1	1	1	1	0	1
2-amino adipic acid	1	1	1	0	1	1
unknown133	1	0	1	1	1	1
unknown43	1	0	1	1	1	1
unknown99	1	1	0	1	1	1
spermidine	1	1	1	1	1	0
unknown21	1	1	0	1	1	0
unknown102	1	1	1	0	0	1
unknown74	1	1	1	1	0	0
fructose2	1	0	1	1	1	0
glucose-6-phosphate2	0	1	1	0	1	1
unknown94	1	0	1	0	1	1
unknown12	1	1	1	0	1	0
unknown134	1	1	0	1	0	1
unknown42	1	1	1	0	0	0
unknown83	1	1	0	0	1	0
ascorbic acid	1	0	1	1	0	0
unknown131	1	0	0	1	0	0
adipic acid	0	1	0	0	1	0
unknown18	1	0	1	0	0	0
unknown10	1	0	0	0	0	1
unknown139	0	0	1	0	0	0
unknown51	0	0	1	0	0	0

Table S3. The 26 metabolites affected by postmortem delay in CBC and PFC.

	CBC	PFC
2-amino adipic acid	yes	no
4-hydroxyproline	yes	no
6-phosphogluconic acid	no	yes
alanine	no	yes
butyric acid	no	yes
citric acid	yes	no
unknown135	yes	no
fructose	yes	yes
fructose2	yes	yes
gaba	yes	no
glucose	yes	yes
glucose-1-phosphate	no	yes
glucose-6-phosphate2	yes	yes
glycerol	yes	no
glycerol-2-phosphate	yes	no
glycerol-3-phosphate	yes	yes
pantothenic acid	yes	no
unknown10	yes	no
unknown126	yes	yes
unknown130	yes	no
unknown23	no	yes
unknown36	yes	no
unknown55	no	yes
unknown62	yes	no
unknown64	yes	no
unknown93	yes	no

Table S4. The age-related metabolites.

The 1-labels indicate metabolites with significant concentration changes with age. The 0-labels indicate metabolites not passing the significance cutoff.

	CBC			PFC		
	chimpanzee	human	macaque	chimpanzee	human	macaque
3-hydroxybutyric acid	0	0	1	0	0	1
octadecenoic acid	0	0	0	0	1	0
decanoic acid	0	0	0	0	1	0
ascorbic acid	1	1	1	1	1	1
aspartic acid	0	1	1	0	1	1
benzoic acid	0	1	1	0	0	0
cholesterol	0	1	1	0	1	1
creatinine	0	1	0	0	1	1
dehydroascorbic acid	0	1	1	1	1	0
erythronic acid	0	0	1	0	0	0
fumaric acid	0	0	0	0	1	0
glutamate	0	1	0	0	1	0
glutamine	0	1	1	0	0	1
glycine	0	1	1	0	1	0
glycolic acid	0	1	1	0	1	1
hexadecanoic acid	0	0	0	0	1	0
histidine	0	1	1	0	1	1
hydroxypyridine	0	0	1	0	0	1
isoleucine	0	1	0	0	1	0
leucine	0	1	1	0	1	0
methionine	0	1	1	0	1	0
nicotinamide	0	0	0	0	1	1
ornithine	0	1	1	0	1	0
oxalic acid	0	0	1	0	0	1
oxoproline	0	1	1	0	1	0
phenylalanine	0	1	1	0	1	0
putrescine	0	1	0	0	1	0
pyruvic acid	0	0	0	0	1	1
ribitol	0	1	0	0	1	0
serine	0	1	1	0	1	0
spermidine	0	0	1	0	0	1
succinic acid	0	1	1	0	0	1
taurine	0	1	1	1	1	0
tetracosanoic acid	0	0	0	0	1	0
tetradecanoic acid	0	0	0	0	1	0
threonic acid	0	1	1	0	1	0
threonine	0	1	1	0	1	0
tryptophan	0	1	1	0	1	0
tyrosine	0	1	1	0	1	0
uridine	0	0	0	0	1	0
valine	0	1	0	0	1	0
unknown1	0	0	1	0	0	0
unknown102	0	1	0	0	0	0

unknown108	0	0	1	0	0	1
unknown12	0	1	1	0	0	0
unknown128	0	1	0	0	0	1
unknown131	0	0	1	0	0	0
unknown132	0	1	1	1	1	1
unknown133	0	1	0	0	1	0
unknown134	0	1	0	0	0	0
unknown136	0	1	0	0	1	0
unknown137	0	1	0	0	1	0
unknown138	0	1	1	0	1	1
unknown18	0	1	1	0	0	0
unknown19	0	0	1	0	0	0
unknown2	0	1	1	0	0	1
unknown20	0	0	0	0	1	0
unknown22	0	0	0	0	1	0
unknown27	0	1	1	0	1	0
unknown3	0	1	0	0	1	0
unknown30	0	0	0	0	1	0
unknown35	0	0	1	0	0	0
unknown4	0	1	1	0	1	0
unknown40	0	0	1	0	0	0
unknown42	0	1	1	0	0	0
unknown43	0	0	1	0	0	0
unknown45	0	0	1	0	1	1
unknown48	1	1	1	1	1	1
unknown49	0	1	0	0	1	0
unknown5	0	0	0	0	1	0
unknown6	0	0	0	0	1	0
unknown66	0	1	0	0	1	1
unknown7	0	0	0	0	0	1
unknown71	0	1	1	0	0	0
unknown74	0	0	1	0	0	1
unknown79	0	1	0	0	0	1
unknown83	0	0	0	0	1	0
unknown85	0	0	0	0	1	0
unknown91	0	1	1	0	0	1
unknown94	0	0	0	0	1	0
unknown99	0	0	1	0	0	0

Table S5. The species-specific metabolites.

	CBC								PFC							
	human			macaque			chimp.		human†			macaque			chimp.	
	F*	SS*	SC	F	SS	SC	SS	SC	F	SS	SC	F	SS	SC	SS	SC
hexadecanoic acid	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0
Octadecenoic acid	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0
unknown128	0	0	0	1	0	0	0	0	1	1	1	0	0	0	0	0
glycolic acid	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
spermidine	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
unknown20	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
unknown6	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
glutamate	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
ribitol	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0
tetracosanoic acid	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
ascorbic acid	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0
Decanoic acid	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
histidine	0	0	0	1	1	1	0	0	1	0	0	0	0	0	0	0
hydroxypyridine	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
oxoproline	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
unknown2	0	0	0	1	1	1	0	0	1	0	0	0	0	0	0	0
unknown22	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
unknown27	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
unknown30	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
unknown4	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0
unknown49	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
unknown5	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
unknown79	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0
unknown83	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
unknown85	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
taurine	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
unknown45	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0
cholesterol	0	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0
3-hydroxybutyric acid	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
aspartic acid	0	0	0	1	1	1	0	0	0	0	0	1	0	0	1	0
benzoic acid	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
dehydroascorbic acid	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
erythronic acid	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
fumaric acid	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
glutamine	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0
glycine	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0
isoleucine	0	0	0	1	1	0	0	0	0	0	0	1	1	1	0	0
leucine	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0
methionine	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0
nicotinamide	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
ornithine	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0

oxalic acid	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
phenylalanine	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
putrescine	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0
unknown138	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0
serine	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
threonic acid	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
threonine	0	0	0	1	1	1	0	0	0	0	0	1	1	1	0	0
tryptophan	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
tyrosine	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
unknown1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
unknown108	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
unknown12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
unknown132	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0
unknown134	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
unknown137	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0
unknown18	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
unknown19	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
unknown21	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
unknown40	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
unknown42	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
unknown48	0	0	0	1	0	0	0	0	0	0	0	1	1	1	0	0
unknown66	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
unknown71	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0
unknown74	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
unknown91	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
valine	0	0	0	1	1	0	0	0	0	0	0	1	1	1	0	0

* F: full dataset; SS: 11 individual subset with stage-of-life approach; SC: 11 individual subset with chronological approach.

† The overlap between 24 and 10 human specific metabolites identified in PFC using full set or 11-individuals subset with stage-of-life approach is 7. To estimate the possibility that the overlap is observed by chance, we randomly choose 10 metabolites from 92 metabolites 1,000 times and calculate the *p*-value which is defined as the frequency of the event in which the overlap between randomly chosen metabolites and 24 human specific is not less than 7. The calculated *p*-value equals 0.005. Accordingly, the overlaps of human specific metabolite in PFC between full set and 11-individual subset with chronological approach or between both subsets are more than expected. There are 6 shared between full set and 11-individual subset with chronological approach with *p*-value equals 0.004 and 4 between both subsets with *p*-value equals 0.001.

Table S6. Summary of the 118 metabolites.

		annotated	hypothetical	total
detected		61	57	118
postmortem		17	9	26
used in analyses		44	48	**92**
age-related in 3 species	CBC*	32	31	63
	PFC*	39	27	66
	Union**	41	40	**81**
human-macaque difference	CBC*	27	16	43
	PFC*	29	20	49
	Union**	36	26	**62**
human-specific ***	CBC*	3	3	6
	PFC*	11	13	24
	Union**	13	16	**29**

* Union of age-related metabolic changes in three species (FDR<10% in each species), analysis based on the full set of individuals

** Union of the two brain regions

*** Based on the full-set of individuals

Table S7. Human metabolites affected by diet or exercise.

Factor	Tissue	Organism	Metabolite	Detected	Category	Reference
diet	urine	human	creatinine	yes	increased in high-meat	(2)
diet	urine	human	taurine	yes		
diet	urine	human	carnitine	no		
diet	urine	human	trimethylamine-N-oxide	no		
diet	urine	human	methylhistidine	no		
diet	urine	human	acetylcarnitine	no		
diet	urine	human	glutamine	yes		
diet	urine	human	p-hydroxyphenylacetate	no	increased in vegetarian	
diet	urine	human	N-acetylglutamate	no		
diet	urine	human	N,N,N trimethyllysine	no	increased in low-meat	
exercise	urine	human	2-hydroxyisovalerate	no	increased in exercise	(7)
exercise	urine	human	2-oxoisocaproate	no		
exercise	urine	human	3-hydroxyisobutyrate	no		
exercise	urine	human	3-methyl-2-oxovalerate	no		
exercise	urine	human	2-oxoisovalerate	no		
exercise	urine	human	2-hydroxybutyrate	no		
exercise	urine	human	lactate	no		
exercise	urine	human	alanine	no		
exercise	urine	human	pyruvate (pyruvic acid)	yes		
exercise	urine	human	2-oxovalerate	no		
exercise	urine	human	inosine	no		
exercise	urine	human	fumarate (fumaric acid)	yes		
exercise	urine	human	hypoxanthine	no		
exercise	urine	human	citrate^	no	decreased in exercise	
exercise	urine	human	trimethylamine N-oxide	no		
exercise	urine	human	taurine	yes		
exercise	urine	human	glycine^	yes		

exercise	urine	human	allantoin	no		
exercise	urine	human	phenylalanine	yes		
exercise	urine	human	hippurate	no		
exercise	urine	human	tryptophan	yes		
exercise	urine	human	formate	no		
exercise	urine	human	leucine	yes	no significant change in exercise	
exercise	urine	human	valine	yes		
exercise	urine	human	isoleucine	yes		
exercise	urine	human	3-hydroxybutyrate (3-hydroxybutyric acid)	yes		
exercise	urine	human	2-hydroxyisobutyrate	no		
exercise	urine	human	acetate	no		
exercise	urine	human	acetoacetate	no		
exercise	urine	human	succinate (succinic acid)*	yes		
exercise	urine	human	dimethylamine	no		
exercise	urine	human	creatinine*	yes		
exercise	urine	human	cis-aconitate	no		
exercise	urine	human	malonate	no		
exercise	urine	human	carnitine*	no		
exercise	urine	human	N-methylnicotinamide	no		
exercise	urine	human	glucuronate	no		
exercise	urine	human	allantoate	no		
exercise	urine	human	trans-aconitate	no		
exercise	urine	human	tyrosine	yes		
exercise	urine	human	histidine	yes		
exercise	urine	human	1-methylhistidine	no		
exercise	urine	human	3-methylhistidine	no		
exercise	muscle	human	glutamate	yes	decreased in exercise	(4)
exercise	muscle	human	phosphocreatine	no		
exercise	muscle	human	glycogen	no		
exercise	muscle	human	carnitine	no		

exercise	muscle	human	alanine	no	increased in exercise	
exercise	muscle	human	creatine	no		
exercise	muscle	human	lactate	no		
exercise	muscle	human	pyruvate (pyruvic acid)	yes		
exercise	muscle	human	malate	no		
exercise	muscle	human	fumarate (fumaric acid)	yes		
exercise	muscle	human	citrate^	no		
exercise	muscle	human	isocitrate	no		
exercise	muscle	human	acetylcarnitine	no		
exercise	muscle	human	ATP	no	no significant change in exercise	
exercise	muscle	human	free CoA	no		
exercise	muscle	human	creatine	no	increased in exercise	(3)
exercise	muscle	human	ATP^	no		
exercise	muscle	human	phosphocreatine*	no	no significant change in exercise	
diet	plasma	mouse	Histamine	no	increased in high-fat food	(1)
diet	plasma	mouse	α-Hydroxyisocaproate	no		
diet	plasma	mouse	Linoleate	no		
diet	plasma	mouse	Arachidonate	no		
diet	plasma	mouse	adenosine 5'-Monophosphate	no		
diet	plasma	mouse	Nicotinamide	yes		
diet	plasma	mouse	α-Tocopherol	no		
diet	plasma	mouse	EDTA	no		
diet	plasma	mouse	Kynurenate	no	decreased in high-fat food	
diet	plasma	mouse	Glutamine	yes		
diet	plasma	mouse	Glutamate	yes		
diet	plasma	mouse	Ornithine	yes		
diet	plasma	mouse	Arginine	no		
diet	plasma	mouse	4-Guanidinobutanoate	no		
diet	plasma	mouse	Asparagine	no		
diet	plasma	mouse	4-Methyl-2-oxopentanoate	no		

diet	plasma	mouse	Glycine	yes		
diet	plasma	mouse	1,5-Anhydroglucitol	no		
diet	plasma	mouse	malate	no		
diet	plasma	mouse	Citrate	no		
diet	plasma	mouse	fumarate (fumaric acid)	yes		
diet	plasma	mouse	Myo-inositol	no		
diet	plasma	mouse	Myristate	no		
diet	plasma	mouse	Acetyl-carnitine	no		
diet	plasma	mouse	Hexanoyl-carnitine	no		
diet	plasma	mouse	Carnitine	no		
diet	plasma	mouse	Glycerol	no		
diet	plasma	mouse	γ-Glutamyl-tyrosine	no		
diet	plasma	mouse	Hippurate	no		
exercise	liver	rat	ATP^	no	decreased in exercise	(6)
exercise	liver	rat	TAN	no		
exercise	liver	rat	AMP	no	increased in exercise	
exercise	liver	rat	Adenosine^	no		
exercise	liver	rat	Hypoxanthine	no		
exercise	liver	rat	Uric acid	no		
exercise	liver	rat	ADP	no	no change in exercise	
exercise	liver	rat	IMP	no		
exercise	liver	rat	Inosine*	no		
exercise	liver	rat	Alanine	no	increased in exercise	
exercise	liver	rat	Glycine^	yes		
exercise	liver	rat	Threonine	yes		
exercise	liver	rat	Cysteine	no		
exercise	liver	rat	Glutamine	yes		
exercise	liver	rat	Ornithine	yes		
exercise	liver	rat	Succinate (succinic acid)	yes		
exercise	liver	rat	fumarate (fumaric acid)	yes		
exercise	liver	rat	Malate	no		

exercise	liver	rat	Dihydroxybutyric acidaba	no		(5)
exercise	liver	rat	Oleic acid	no		
exercise	liver	rat	Linoleic acid	no		
exercise	liver	rat	Stearic acid	no		
exercise	liver	rat	Ribofuranose	no		
exercise	liver	rat	α-Ribopyranose	no		
exercise	liver	rat	Gluconic acid	no		
exercise	liver	rat	Hypoxanthine	no		
exercise	liver	rat	Xanthine	no		
exercise	liver	rat	Phosphoric acid	no		
exercise	liver	rat	Glycerol phosphate	no		
exercise	liver	rat	Glucose phosphate	no		
exercise	liver	rat	Urea	no		
exercise	liver	rat	β-Aminoisobutyric acid	no		
exercise	liver	rat	Aminomalonic acid	no		
exercise	liver	rat	Creatinine	yes		
exercise	liver	rat	Pentanoic acid	no		
exercise	liver	rat	Ascorbic acid	yes		
exercise	liver	rat	Stearic acid glycerol	no	decreased in exercise	
exercise	liver	rat	α-Glucose	no		
exercise	liver	rat	Glucitol	no		
exercise	liver	rat	β-Glucose	no		
exercise	liver	rat	Saccharide 2	no		
exercise	liver	rat	Saccharide 3	no		
exercise	liver	rat	Saccharide 4	no		
exercise	liver	rat	Uridine	yes		
exercise	liver	rat	Adenosine^	no		
exercise	liver	rat	Diglycerol phosphate	no		
exercise	liver	rat	2-Hydroxybutyric acid	no	changed in	
exercise	liver	rat	Lactate	no		
exercise	liver	rat	Arachidonic acidbca	no		

exercise	liver	rat	β-Ribopyranose	no	exercise#	
exercise	liver	rat	Saccharide 1	no		
exercise	liver	rat	Uric acid	no		
exercise	liver	rat	Valine	yes	no change in exercise	
exercise	liver	rat	Leucine	yes		
exercise	liver	rat	Isoleucine	yes		
exercise	liver	rat	Serine	yes		
exercise	liver	rat	Methionine	yes		
exercise	liver	rat	Aspartic acid	yes		
exercise	liver	rat	Tyrosine	yes		
exercise	liver	rat	Palmitic acid glycerol	no		
exercise	liver	rat	Niacinamide	no		
exercise	liver	rat	Inositol	no		
exercise	liver	rat	Cholesterol	yes		
* There are 6 metabolites (succinic acid, creatinine, carnitine, ATP, phosphocreatine, and inosine) that identified as changed in exercise in some studies but not changed in other studies are considered as changed in exercise in this study.						
# There are 6 metabolites (2-Hydroxybutyric acid, Lactate, Arachidonic acidbca, β-Ribopyranose, Saccharide 1, Uric acid) showing different response in exhaustive exercise and endurance training compared with controls are considered as changed in exercise in this study.						
^ There are 4 metabolites (citrate, ATP, adenosine, glycine) showing opposite responses to exercise among different studies will also considered as changed in corresponding factor in this study.						

Table S8. Overlap between the species-specific metabolites and the diet/exercise affected metabolites.

Brain region	Species specific		Affected by diet^	Affected by exercise	Not affected by exercise
	Species	Number *			
PFC	human	24 (11)	*glutamate#*	glutamate#	histidine
	macaque	20 (15)	*ornithine*; *glycine*; *fumaric acid*; *nicotinamide*	fumaric acid; glycine; phenylalanine; *threonine*; *ornithine*; *ascorbic acid*	leucine; valine; isoleucine; *methionine*; *aspartic acid*
CBC	human	6 (3)	taurine	taurine; *ascorbic acid*	-
	macaque	33 (20)	glutamine; *ornithine*;	glycine; phenylalanine;	leucine; valine; isoleucine;

			ornithine	tryptophan; *threonine*; *glutamine*; *ornithine*	tyrosine; histidine; *serine*; *methionine*; *aspartic acid*

* Numbers in parentheses were the amount of metabolites with known annotation;

^ Metabolites in italic were those detected as affected by diet, exercise, or not effected by exercisebased on the data from rat or mouse.

Glutamate has been shown to decreased in high-fat diet in mouse plasma (1) and decrease in concentration with increased exercise (4).

Table S9. Types of metabolic differences among species.

Region	Differential metabolites		Potential cause		
	Type	Number	Life-span	Mean	Pattern
PFC	human-specific	24	0	17	7
	macaque-specific	20	1	11	8
	rest	5	4	0	1
	total	49	5 (10%)	28 (57%)	16 (32%)
CBC	human-specific	6	1	2	3
	macaque-specific	33	2	26	5
	rest	4	2	2	0
	total	43	5 (12%)	30 (70%)	8 (19%)

References

1. Li LO*, et al.* (2010) Early hepatic insulin resistance in mice: a metabolomics analysis. *Mol Endocrinol* 24:657-666.
2. Stella C*, et al.* (2006) Susceptibility of human metabolic phenotypes to dietary modulation. *J Proteome Res* 5:2780-2788.
3. Burgomaster KA, Heigenhauser GJ,Gibala MJ (2006) Effect of short-term sprint interval training on human skeletal muscle carbohydrate metabolism during exercise and time-trial performance. *J Appl Physiol* 100:2041-2047.
4. Howarth KR, LeBlanc PJ, Heigenhauser GJ,Gibala MJ (2004) Effect of endurance training on muscle TCA cycle metabolism during exercise in humans. *J Appl Physiol* 97:579-584.
5. Huang CC, Lin WT, Hsu FL, Tsai PW,Hou CC (2010) Metabolomics investigation of exercise-modulated changes in metabolism in rat liver after exhaustive and endurance exercises. *Eur J Appl Physiol* 108:557-566.
6. Mikami T,Kitagawa J (2006) Intense exercise induces the degradation of adenine nucleotide and purine nucleotide synthesis via de novo pathway in the rat liver. *Eur J Appl Physiol* 96:543-550.
7. Pechlivanis A*, et al.* (2010) (1)H NMR-Based Metabonomic Investigation of the Effect of Two Different Exercise Sessions on the Metabolic Fingerprint of Human Urine. *J Proteome Res* DOI: 10.1021/pr100684t.
8. Antinucci F (1989) *Systematic comparison of early sensorimotor development, in Cognitive structure and development in nonhuman primates, ed Antinucci F* (Erlbaum, Hillsdale (New Jersey)) pp 67–85.
9. Dienske H (1986) *A comparative approach to the question of why human infants develop so slowly, in Primate ontogeny, cognition, and social behaviour, eds Else JG and Lee PC* (Cambridge Univ. Press, New York) pp 147-154.
10. Ferrari PF*, et al.* (2006) Neonatal imitation in rhesus macaques. *PLoS Biol* 4:e302.
11. Watts E (1990) *Evolutionary trends in primate growth and development, in Primate life history and evolution, ed DeRousseau CJ* (Wiley-Liss, New York) pp 89–104.
12. Ben Shaul DM (1963) *The composition of the milk of wild animals, in The International Zoo Year Book, eds Jarvis C and Morris D* (Hutchinson and Company Ltd., London,) pp 333-342.
13. Rosenberg K,Trevathan W (2002) Birth, obstetrics and human evolution. *Bjog* 109:1199-1206.
14. Somel M*, et al.* (2008) Human and chimpanzee gene expression differences replicated in mice fed different diets. *PLoS One* 3:e1504.
15. Lisec J, Schauer N, Kopka J, Willmitzer L,Fernie AR (2006) Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat Protoc* 1:387-396.
16. Fu X*, et al.* (2009) Estimating accuracy of RNA-Seq and microarrays with proteomics. *BMC Genomics* 10:161.
17. Somel M*, et al.* (2010) MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res* 20:1207-1218.
18. Boedigheimer MJ*, et al.* (2008) Sources of variation in baseline gene expression levels from toxicogenomics study control animals across multiple laboratories. *BMC Genomics* 9:285.

19. Li J , Bushel PR, Chu T-M,Wolfinger RD (2009) *Principal Variance Components Analysis: Estimating Batch Effects in Microarray Gene Expression Data, in Batch Effects and Noise in Microarray Experiments: Sources and Solutions, ed Scherer A* (John Wiley & Sons, Ltd, Chichester) doi: 10.1002/9780470685983.ch12.
20. Somel M*, et al.* (2009) Transcriptional neoteny in the human brain. *Proc Natl Acad Sci U S A* 106:5743-5748.
21. Faraway JJ (2002) Practical Regression and ANOVA Using R. Available at: http://cran.r-project.org/doc/contrib/Faraway-PRA.pdf.
22. de Magalhaes JP,Costa J (2009) A database of vertebrate longevity records and their relation to other life-history traits. *J Evol Biol* 22:1770-1774.
23. Storey JD,Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 100:9440-9445.