Supplemental Material

"Cell type-specific analysis by single-cell profiling identifies a stable mammalian tRNA-mRNA interface and increased translation efficiency in neurons"

Gao, Gallardo-Dodd, Kutter

SUPPLEMENTARY FIGURES

SUPPLEMENTARY TABLES

ARRAYEXPRESS ACCESSION

All data analyzed were from publicly available sources that were previously deposited. Please refer to Table S1 for a list of these datasets and their accessions.

CODE ACCESSIBILITY

The main code for analysis of preprocessed data is here: https://github.com/wgao688/sc_tRNA_mRNA. Instructions and code for processing of raw data is available in the Supplemental Code section of the Supplemental Material.

A

	scRNA-seq (Schaum et al. 2018)	scATAC-seq (Cusanovich, Hill et al. 2018)	scRNA-seq (Cao et al. 2020)	scATAC-seq (Domcke et al. 2020)
# tissues	20	13	15	15
# total cell types	110	252	172	126
# cell types after filtering	110	36	172	89
# cell types for tTE analysis	31	31	85	85

B

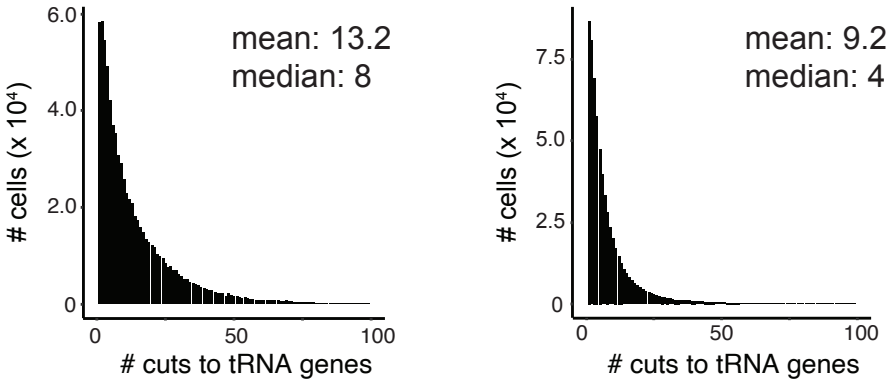

Figure S1. Description of single-cell datasets and scATAC-seq filtering.
(A) Information for scRNA-seq and scATAC-seq datasets for adult mouse and fetal human. The same class of cells (e.g., endothelial cells) but from different tissues were considered distinct cell types. Cell types with insufficient (less than 5,000) scATAC-seq cuts were removed because of unreliable tRNA quantification. After filtering, theoretical translation efficiency (tTE) was calculated only for cell types that were present in both scRNA-seq and scATAC-seq datasets. **(B)** Histograms showing the number of cuts to tRNA genes per cell, for the mouse and human scATAC-seq datasets. Left: mouse, right: human.

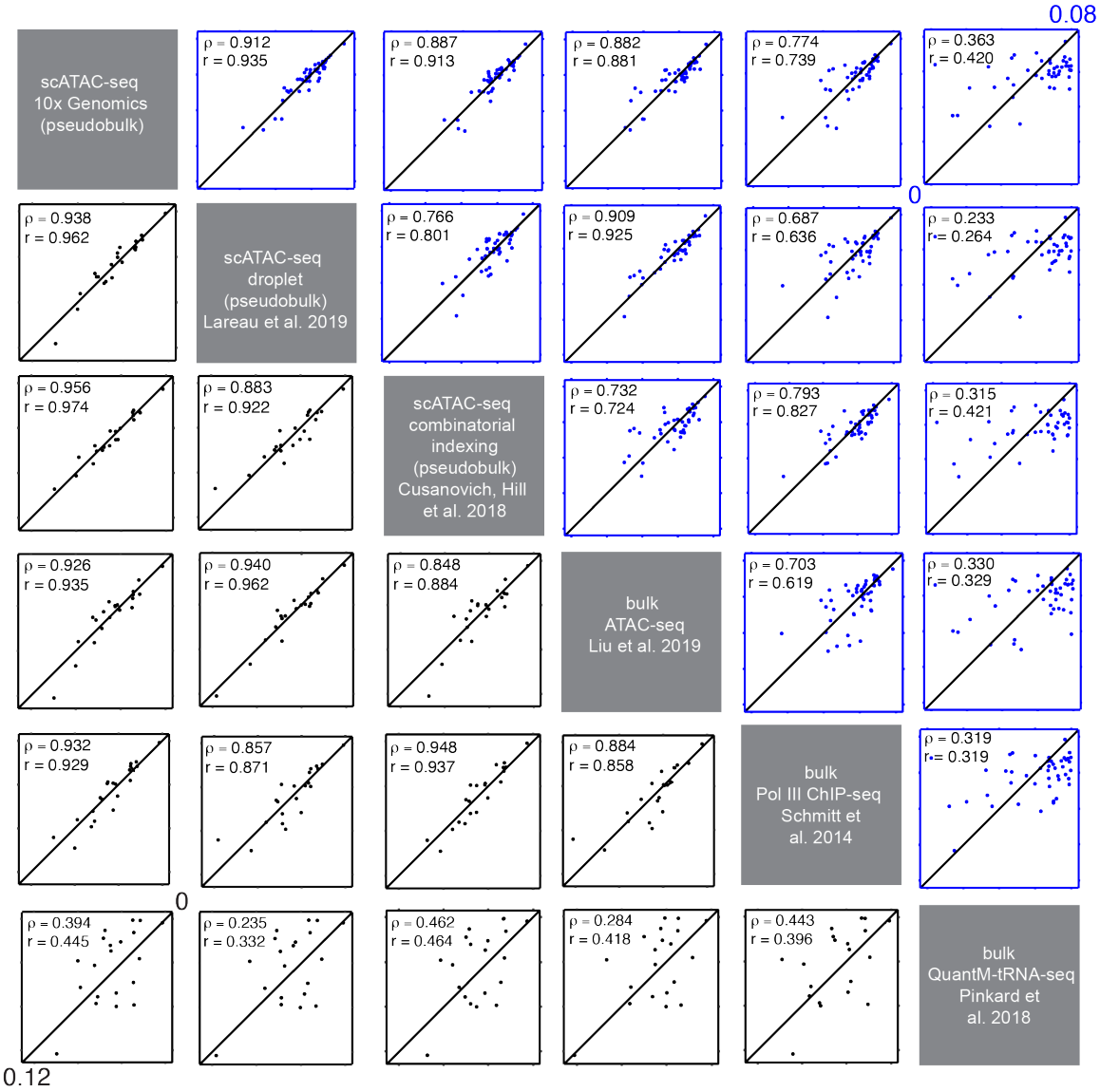

Figure S2. Correlation of tRNA quantifications across mouse brain datasets. Spearman's rank (ρ) and Pearson (r) correlation coefficients of anticodon isoacceptor usage (blue) and amino acid isotype supply (black) across five datasets are shown: 3 scATAC-seq datasets (aggregated in pseudobulk), a bulk ATAC-seq dataset, a Pol III ChIP-seq dataset, and an RNA-based QuantM-tRNA-seq dataset. All datasets were generated using adult mouse brain tissue. Scales, representing fraction of total usage, are indicated on the bottom left and top right corners.

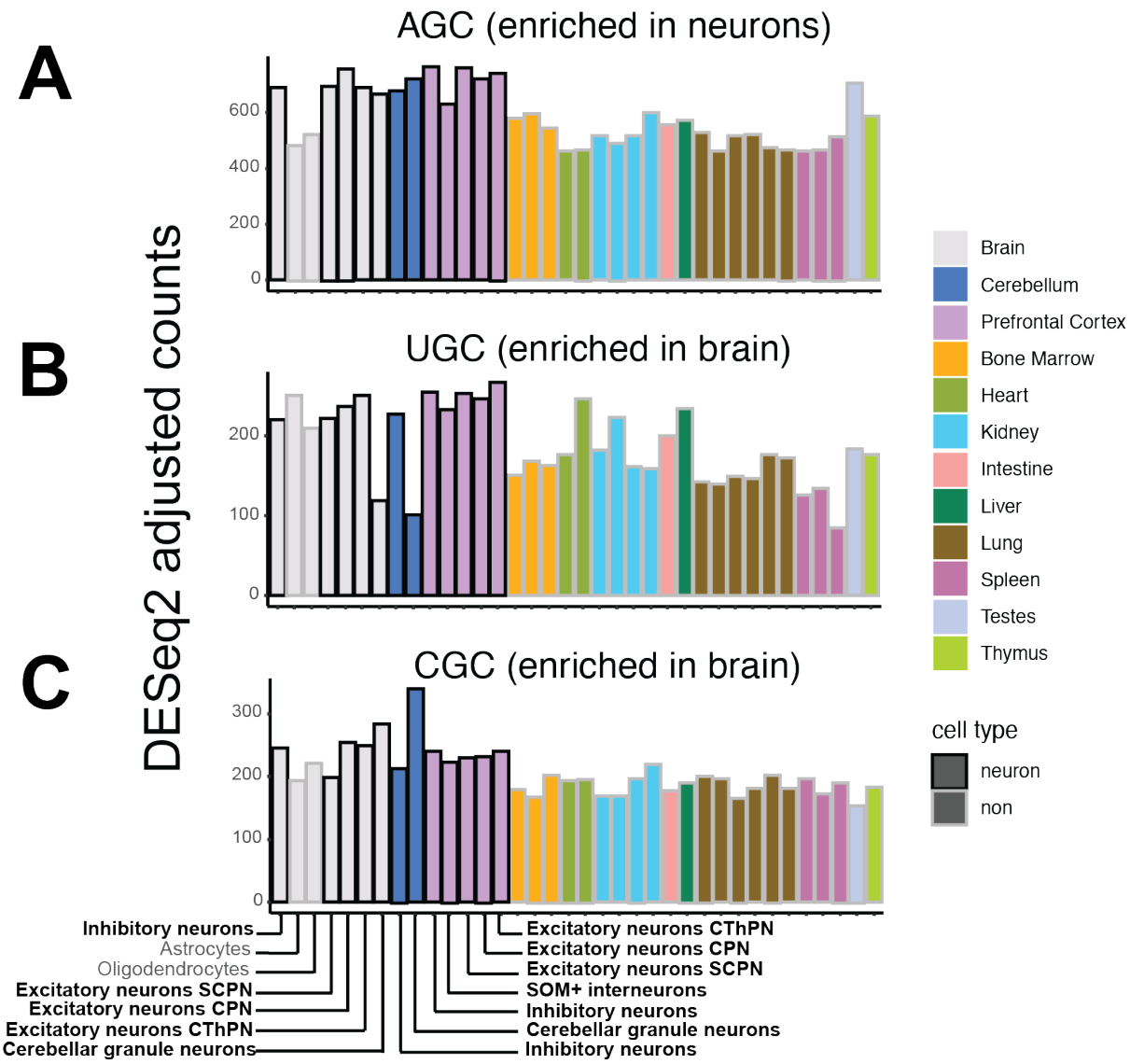

Figure S3. Enrichment of alanine anticodons in mouse brain cells.
(A-C) Bar plots show Ala anticodon usage across cell types. (A) Ala-AGC is enriched in neurons. Glial cells have similar Ala-AGC usage to other cell types. (B) Ala-UGC and (C) Ala-CGC are enriched in both neurons and glial cells, when compared to other cell types.

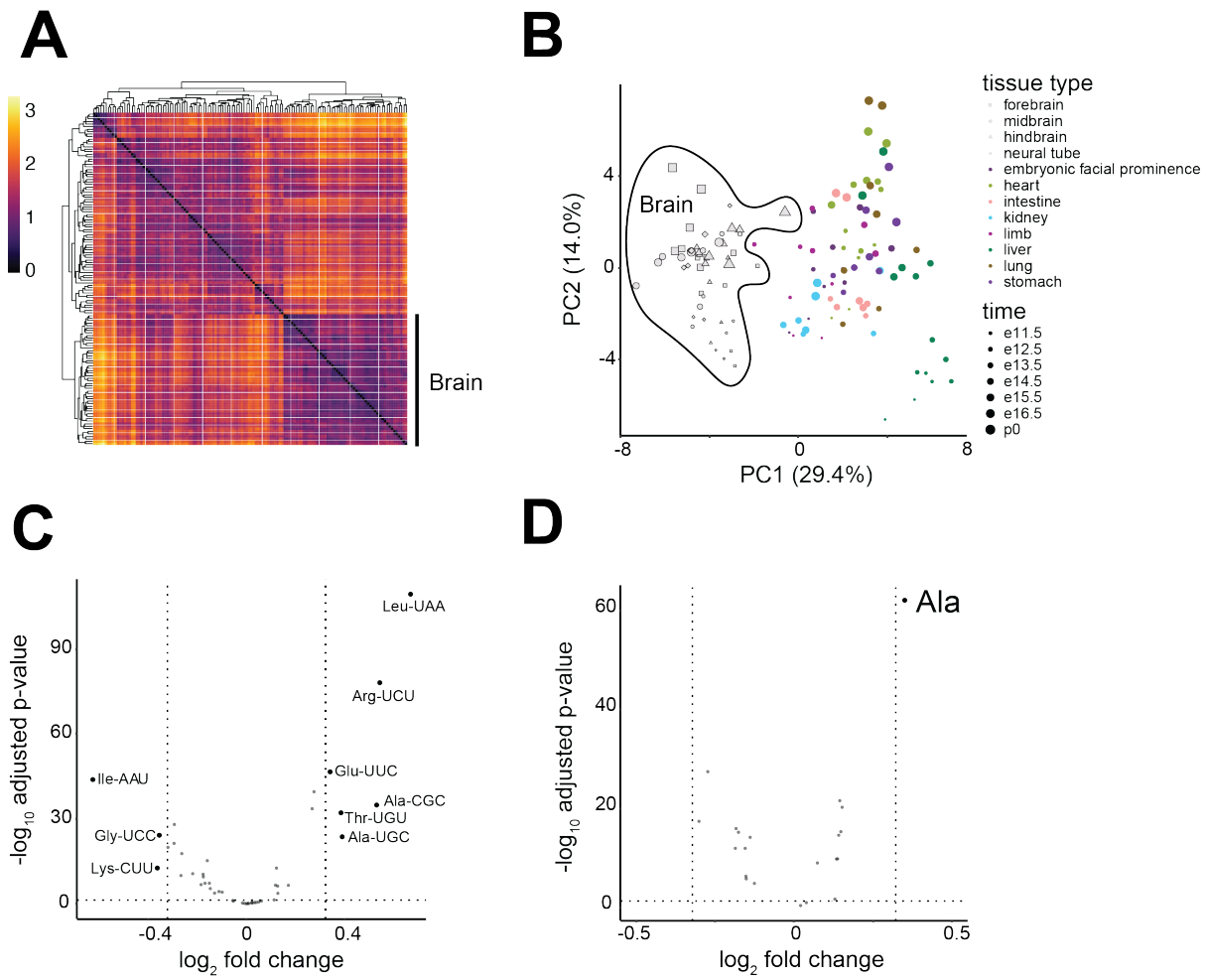

Figure S4. Bulk mouse embryonic ATAC-seq dataset shows brain-specific clustering and enrichment of alanine supply.

(A) Heatmaps illustrates the Euclidean distance in anticodon usage across tissue samples from a bulk ATAC-seq dataset across early mouse development (Gorkin et al. 2020). **(B)** PCA plot separates anticodon usage across tissues. Brain sample cluster (consisting of forebrain, midbrain, hindbrain, and neural tube) is marked. **(C-D)** Volcano plots display differences in (C) anticodon usage and (D) amino acid supply between brain and other samples ($-\log_{10}$ adjusted p-values and $\log_2$ fold change (FC) determined using DESeq2). Vertical lines indicate a fold change greater than 25%.

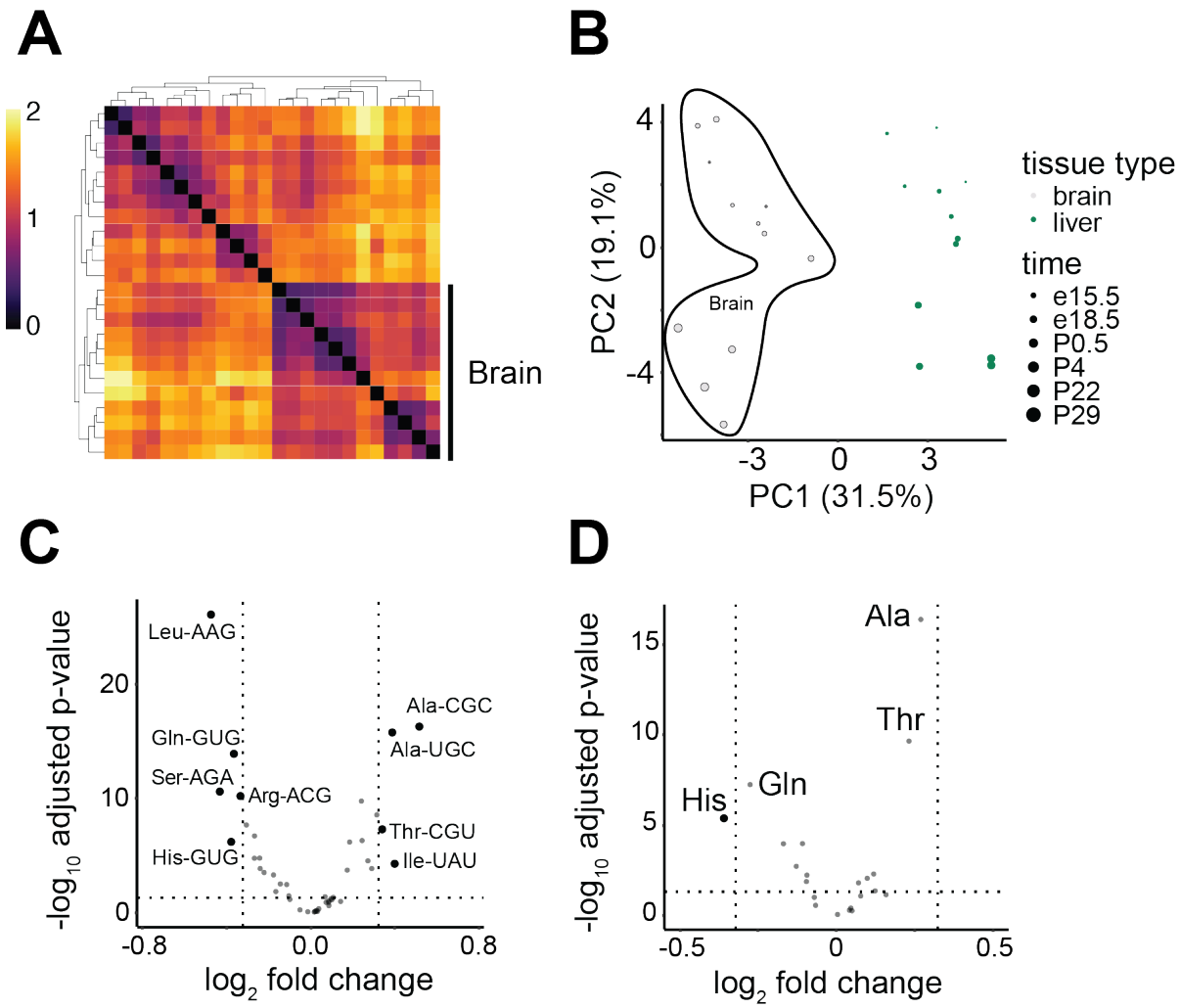

Figure S5. Bulk adult mouse Pol III ChIP-seq dataset shows enrichment of alanine supply in brain.
(A) Heatmaps display the Euclidean distance in anticodon usage of Pol III ChIP-seq samples of mouse liver and brain across development (Schmitt et al. 2014). **(B)** PCA plot separate anticodon usage across brain (highlighted, left) and liver tissues. **(C-D)** Volcano plots reveal significant differences in (C) anticodon usage and (D) amino acid supply between brain and liver samples ($-\log_{10}$ adjusted p-values and $\log_2$ fold change (FC) as determined using DESeq2). Vertical lines indicate a fold change greater than 25%.

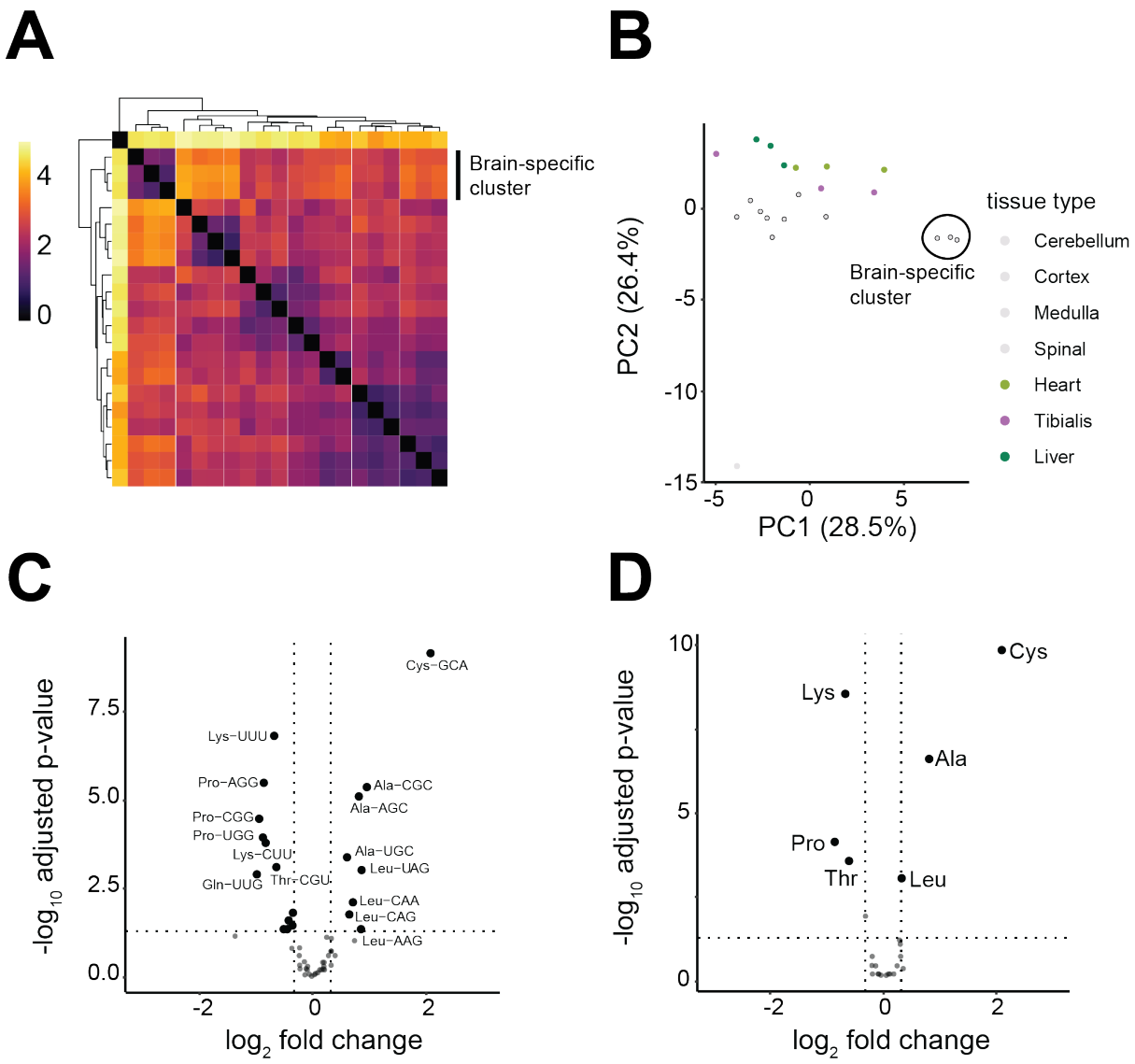

Figure S6. Bulk adult mouse QuantM-tRNA-seq dataset shows brain-specific clustering and enrichment of alanine supply.

(A) Heatmaps show the Euclidean distance in anticodon usage across tissue samples of adult mouse using a bulk QuantM-tRNA-seq dataset (Pinkard et al. 2020). Bar (right) highlights the brain-specific cluster. **(B)** PCA plot separates anticodon usage across tissues. The separately clustering brain samples are highlighted. **(C-D)** Volcano plots show (C) anticodon usage and (D) amino acid supply between brain-specific cluster and other samples ($-\log_{10}$ adjusted p-values and $\log_2$ fold change (FC) as determined using DESeq2). Vertical lines indicate a fold change greater than 25%.

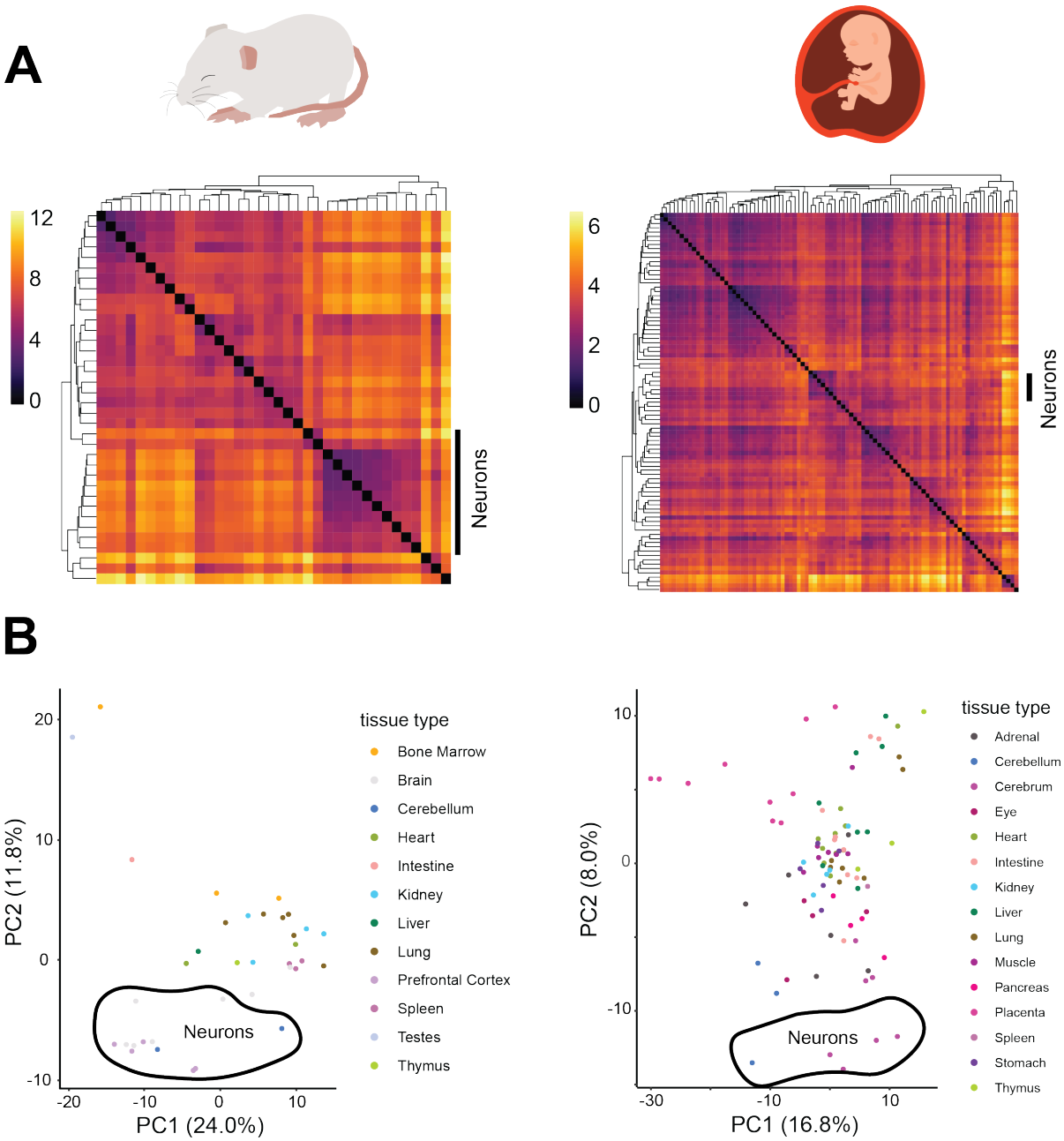

Figure S7. Brain neurons cluster separately in individual tRNA gene usage.
(A) Heatmaps illustrate the Euclidean distance in individual tRNA gene usage across cell types. Bars (right) emphasize neuronal cell type. **(B)** PCA plot separates tRNA gene usage across cell types. The brain neuron specific cluster is indicated. Left: mouse, right: human.

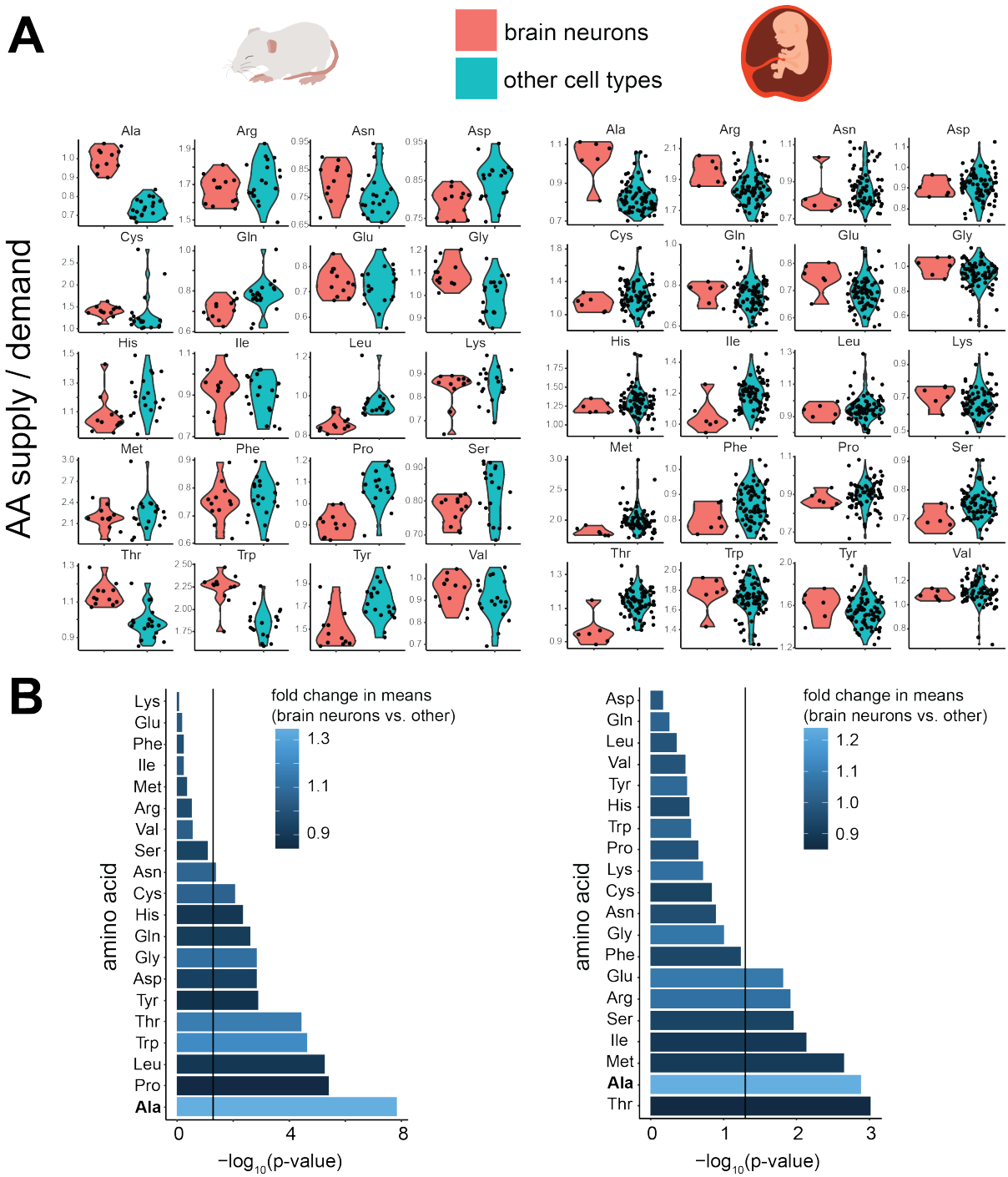

Figure S8. Alanine is the only AA with a statistically significant enrichment in AA supply-to-demand in both mouse and human.
(A) Violin plots show ratios of AA supply-to-demand when comparing brain neurons (red) to all other cell types (green) for the 20 classical AAs. **(B)** Bar plot displays the p-value (Mann-Whitney *U* test) indicating significant differences in AA supply-to-demand ratio for each AA between neurons and other cell types. The vertical line corresponds to p-value of $-\log_{10}(0.05)$. Bars are colored by fold change in mean AA supply-to-demand ratio in brain neurons compared to other cell types. Left: mouse, right: human.

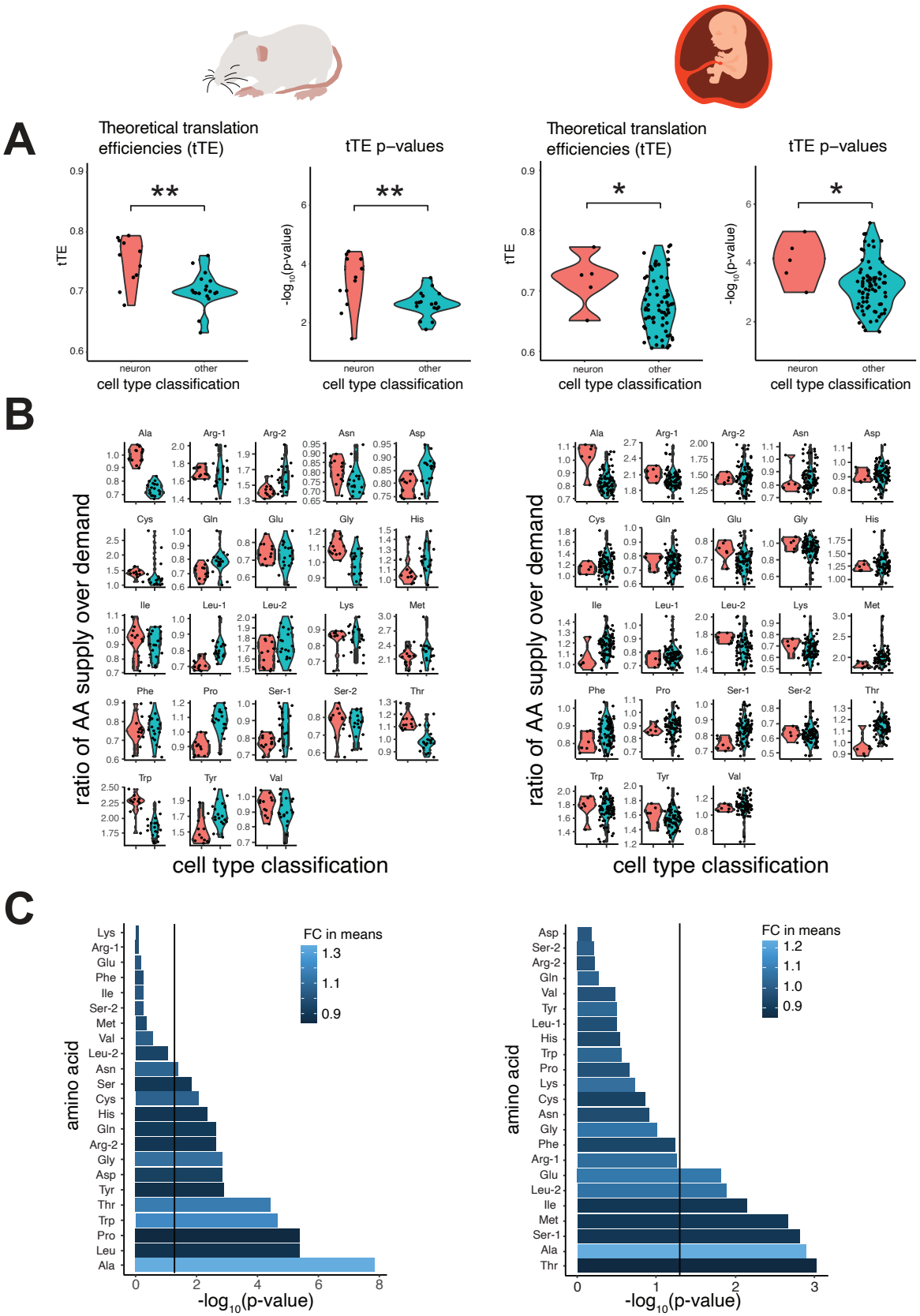

Figure S9. Dividing six box amino acids into two anticodon-codon sets results in similar conclusions

(A) Violin plots display theoretical translation efficiencies (tTEs) and tTE p-values between brain neurons (red) and other cell types (green). **(B)** Violin plots show ratios of AA supply-

to-demand when comparing brain neurons to all other cell types for the 23 groups. Six box amino acids (Arg, Leu, Ser) were divided into two groups based on the anticodons that can complementarily base pair with them, yielding 23 groups instead of 20. Specifically, NCG anticodons can base pair with CGN codons (referred here as Arg-1), while YCT anticodons can base pair with AGR codons (Arg-2). NAG anticodons can base pair with CTN codons (Leu-1) and YAA anticodons can base pair with TTR codons (Leu-2). NGA anticodons can base pair with TCN codons (Ser-1) and RCT anticodons can base pair with AGY codons (Ser-2). N = A, C, G, U; R = A, G; Y = C, U. **(C)** Bar plot illustrates the p-value (Mann-Whitney *U* test) for the difference in AA supply-to-demand ratio for each AA between brain neurons. The vertical line corresponds to p-value $-\log_{10}(0.05)$. Bars are colored by fold change in mean AA supply-to-demand ratio in brain neurons compared to other cell types. Left: mouse, right: human. Asterisks display degree of significance *p<0.05, **p<0.01 (Mann-Whitney *U* test).