Supplemental Material

Schmitt, Rudolph, Karagianni, Fonseca, White, Talianidis, Odom, Marioni, Kutter

"High-resolution mapping of transcriptional dynamics across tissue development reveals a stable mRNA–tRNA interface"

SUPPLEMENTARY FIGURES

Figure S1: Workflow of transcriptome-wide identification and analysis of protein-coding and tRNA genes
Figure S2: Replicate correlations of RNA-seq data during mouse liver and brain development
Figure S3: Replicate correlations of Pol III ChIP-seq data during mouse liver and brain development
Figure S4: Correlation of RNA-seq and Pol III ChIP-seq data during mouse liver and brain development
Figure S5: Many tRNA genes are highly expressed during mouse development
Figure S6: Early developmental stage-specific tRNA genes are lowly expressed
Figure S7: Observed codon usage in mRNA transcriptomes of developing mouse liver
Figure S8: Observed and simulated amino acid and isotype usage in transcriptomes across mouse liver development
Figure S9: mRNA codon usage and Pol III occupancy of tRNA isotypes in developing mouse brain tissue
Figure S10: Highly versus lowly expressed protein-coding genes show no differential codon usage
Figure S11: Observed anticodon usage in tRNA transcriptomes of developing mouse liver
Figure S12: Transcriptomic mRNA codon usage and Pol III binding to tRNA isoacceptors correlate in developing mouse liver and brain
Figure S13. Transcriptomic mRNA codon usage and wobble corrected Pol III binding to tRNA isoacceptors correlate in developing mouse liver and brain.
Figure S14: Transcriptomic mRNA amino acid usage and Pol III binding to tRNA isotypes correlate in developing mouse liver and brain
Figure S15: Differentially expressed tRNA genes show no co-localization with differentially expressed protein-coding genes

SUPPLEMENTARY TABLES

Table S1: Expression values of protein-coding genes per developmental stage (liver)
Table S2: Expression values of protein-coding genes per developmental stage (brain)
Table S3: Differentially expressed protein-coding genes in liver
Table S4: Differentially expressed protein-coding genes in brain
Table S5: GO analysis of differentially expressed protein-coding genes in liver
Table S6: GO analysis of differentially expressed protein-coding genes in brain
Table S7: Expression values of tRNA genes, isoacceptors and isotypes (liver and brain)
Table S8: Differentially expressed tRNA genes in liver
Table S9: Differentially expressed tRNA genes in brain
Table S10: Motif analysis for differentially expressed tRNA genes
Table S11: *p*-values for chromatin accessibility in the vicinity of tRNA genes
Table S12: *p*-values for compensation analysis

SUPPLEMENTARY MATERIAL

SourceCode.zip
SourceCode_dataprocessing.zip
SupplementalData.zip

ARRAYEXPRESS ACCESSION

Pol III ChIP-seq: E-MTAB-2326
RNA-seq: E-MTAB-2328

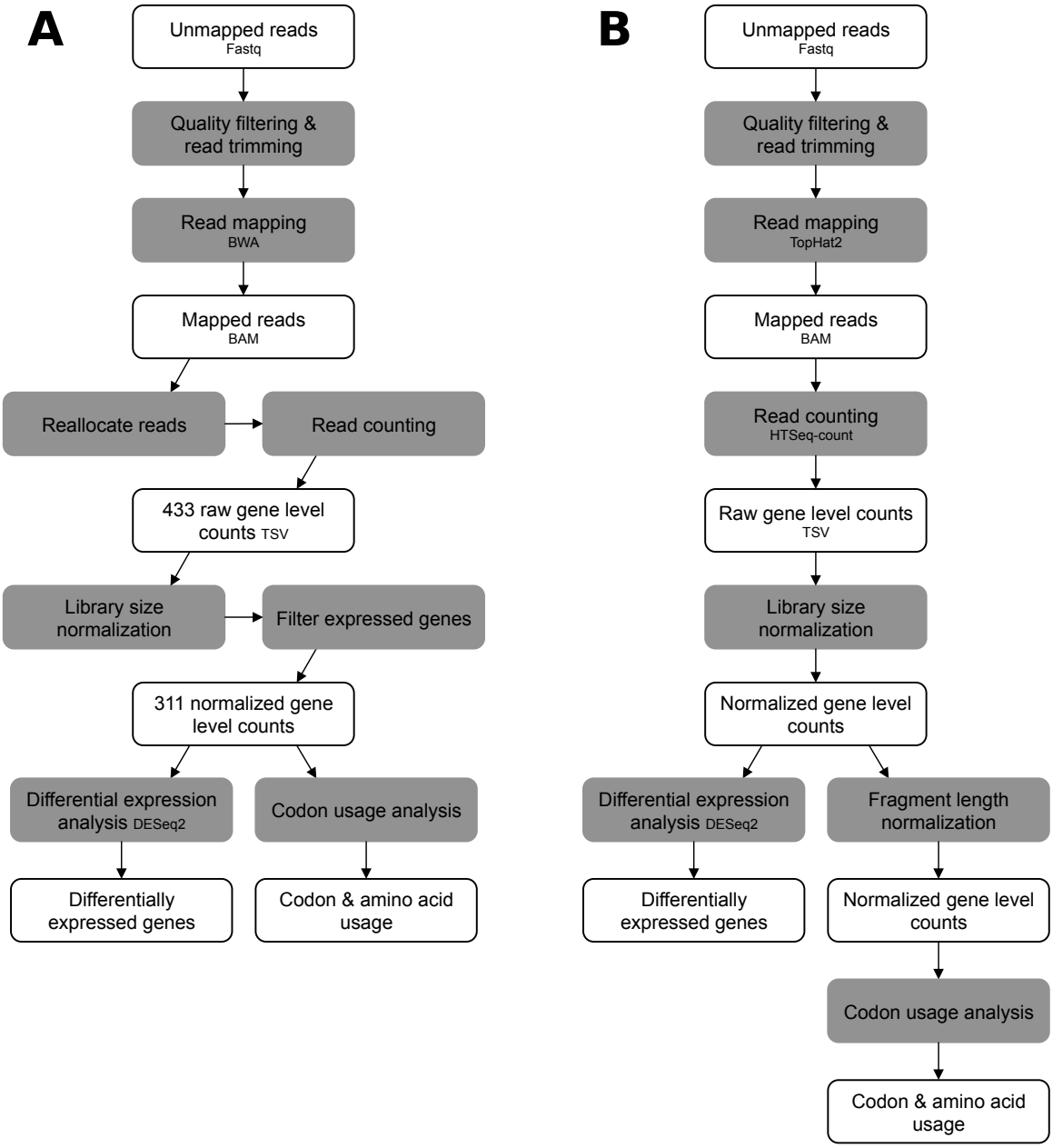

Figure S1. Workflow of transcriptome-wide identification and analysis of protein-coding and tRNA genes

(**A**) RNA-seq analysis of protein-coding gene expression, differential expression analysis and codon usage analysis. (**B**) ChIP-seq analysis of Pol III occupancy at tRNA gene loci, differential expression analysis of tRNA genes, and anticodon isoacceptor abundance.

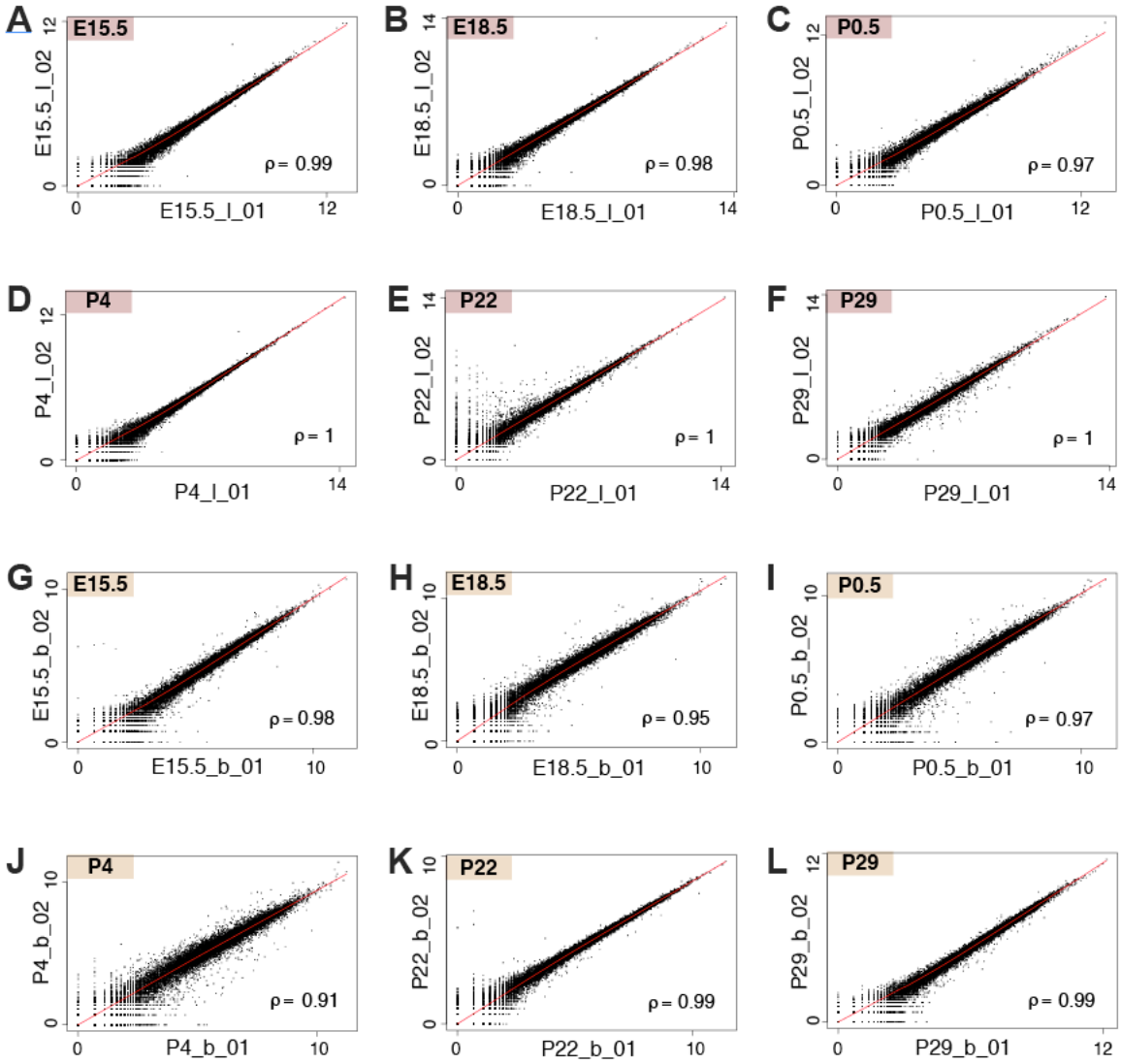

Figure S2: Replicate correlations of RNA-seq data during mouse liver and brain development

Plots present the correlation of protein-coding gene expression between two biological replicates during mouse (**A**–**F**) liver and (**G**–**L**) brain development. Spearman's rank correlation coefficients (ρ) are reported in bottom right of each panel.

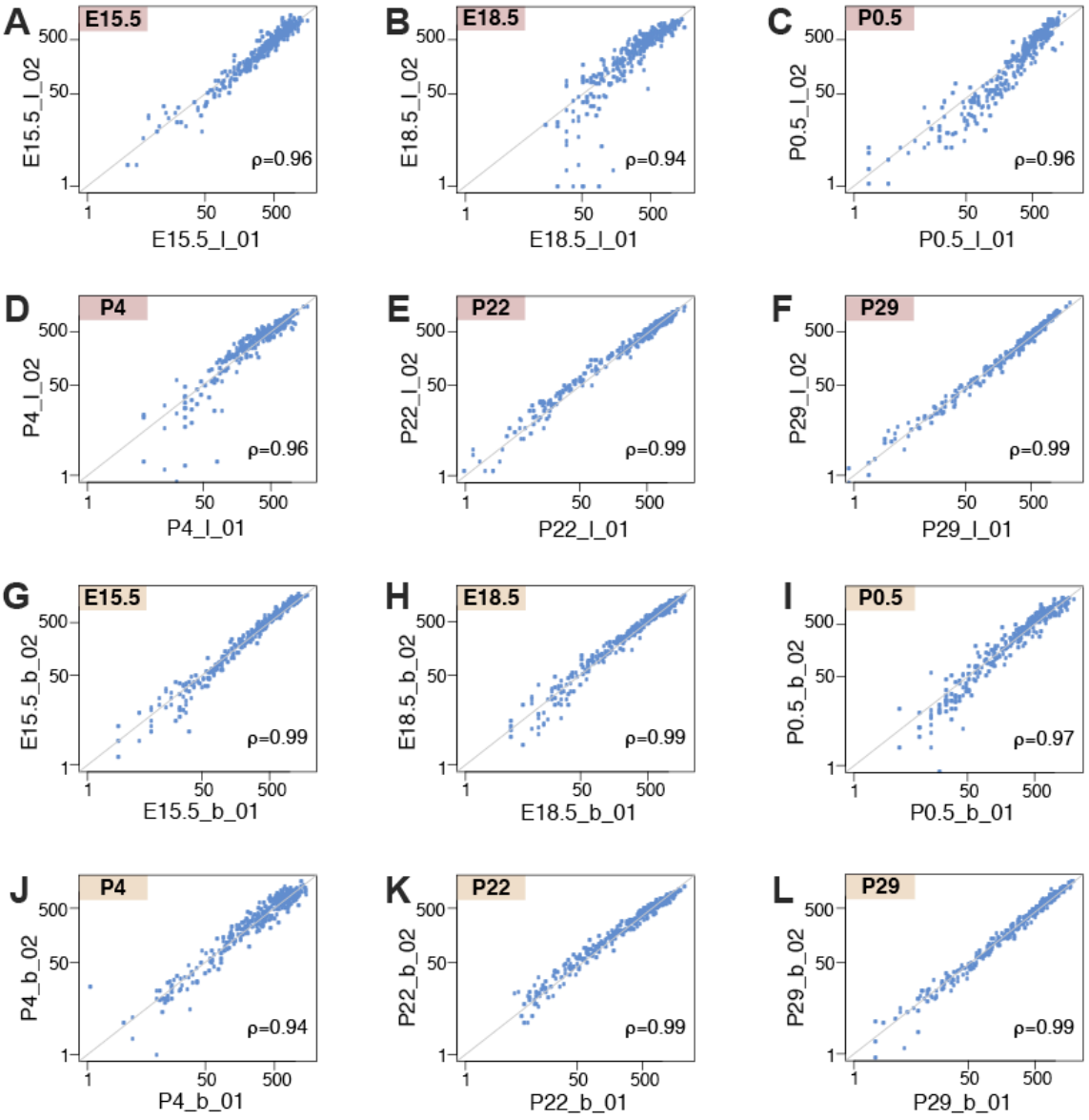

Figure S3: Replicate correlations of Pol III ChIP-seq data during mouse liver and brain development

Plots present the correlation of Pol III binding intensities (log-scaled) to tRNA genes between two biological replicates during mouse (**A**–**F**) liver and (**G**–**L**) brain development. Spearman's rank correlation coefficients (ρ) are reported in bottom right of each panel.

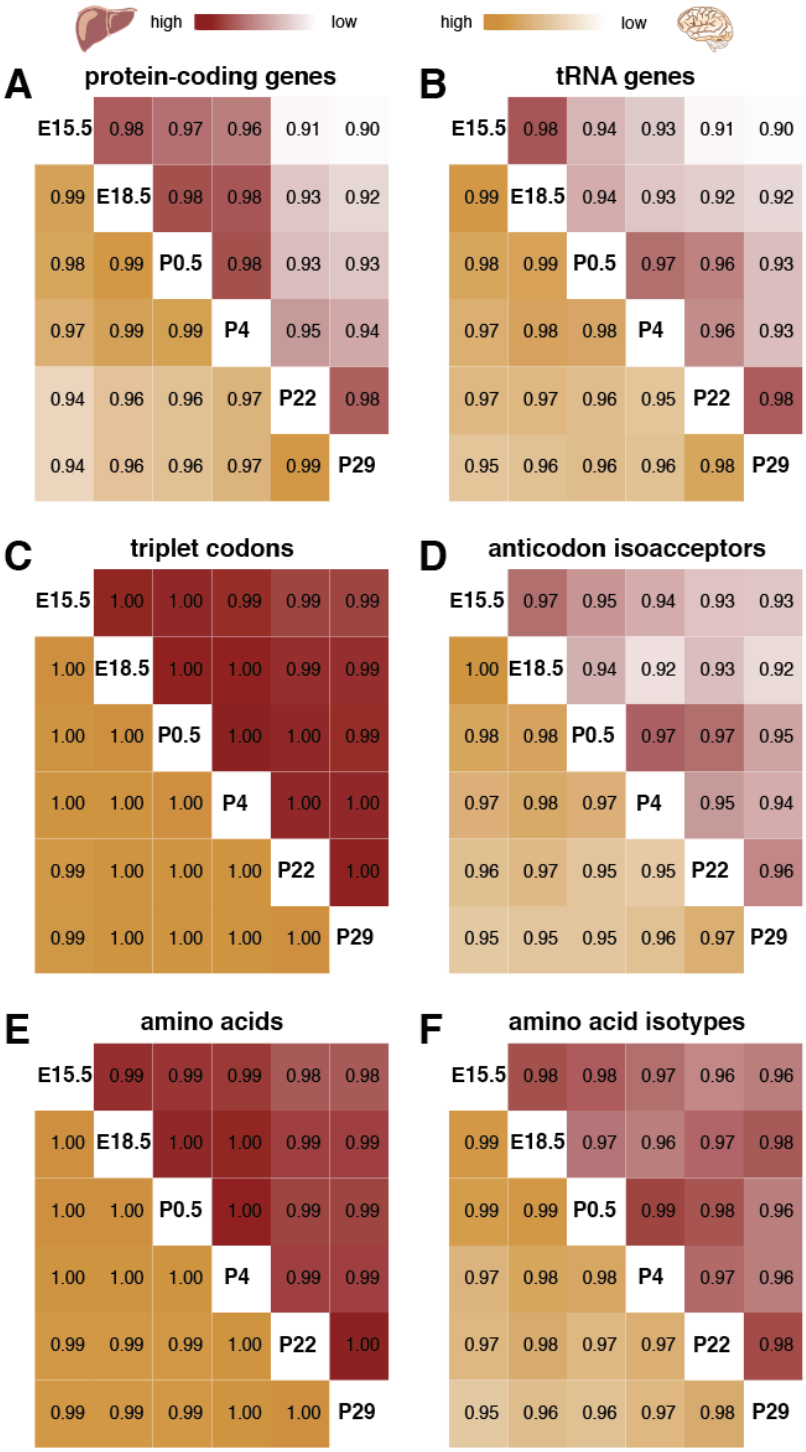

Figure S4. Correlation of RNA-seq and Pol III ChIP-seq data during mouse liver and brain development.
Correlation of (**A**) protein-coding gene expression across developmental stages, (**B**) tRNA gene expression as measured by Pol III occupancy, (**C**) triplet codon usage in protein-coding genes, (**D**) tRNA anticodon isoacceptor, (**E**) amino acid usage of protein-coding genes and (**F**) tRNA amino acid isotype.

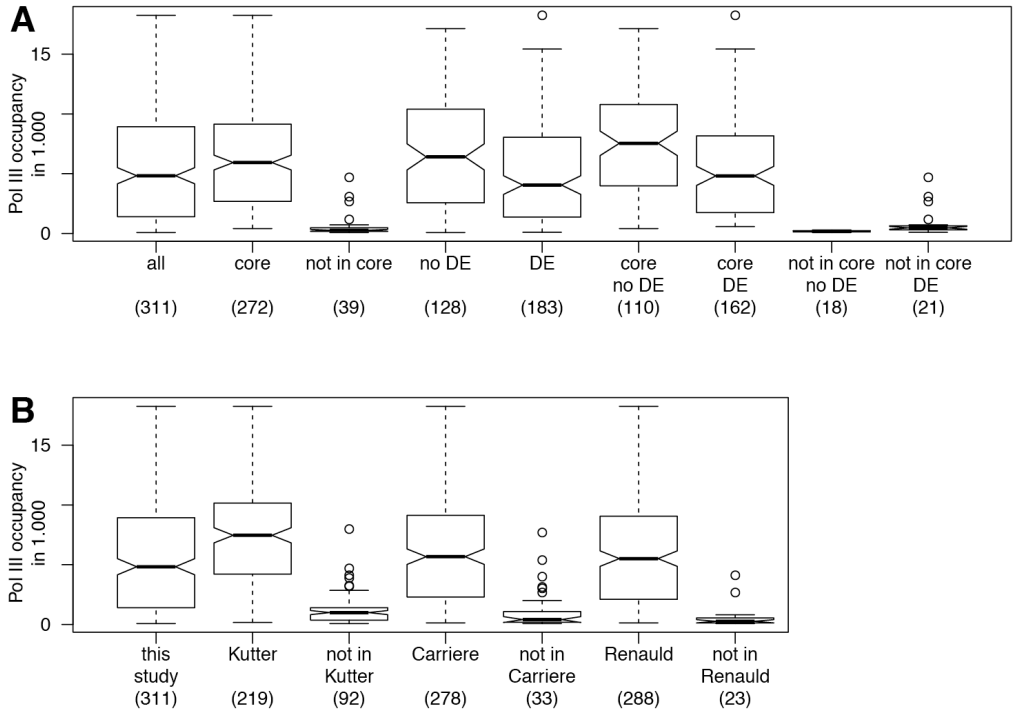

Figure S5. Many tRNA genes are highly expressed during mouse development
Boxplots represents normalized occupancy of Pol III at tRNA genes identified and classified in (**A**) this and (**B**) other studies. In parentheses are the numbers of 311 tRNA genes present in 12 developmental stages according to Figure 3A ("all" or "this study"), which are subdivided by occurrence in all 12 developmental stages ("core") or only present in one up to 11 developmental stages ("not in core"); with ("DE") or without ("no DE") differential expression, overlapping with tRNA genes occupied or unoccupied by Pol III in Kutter et al. (2011) ("Kutter" and "not in Kutter"), Carriere et al. (2012) ("Carriere" and "not in Carriere") and Renauld et al. (2014) ("Renauld" and "not in Renauld")..

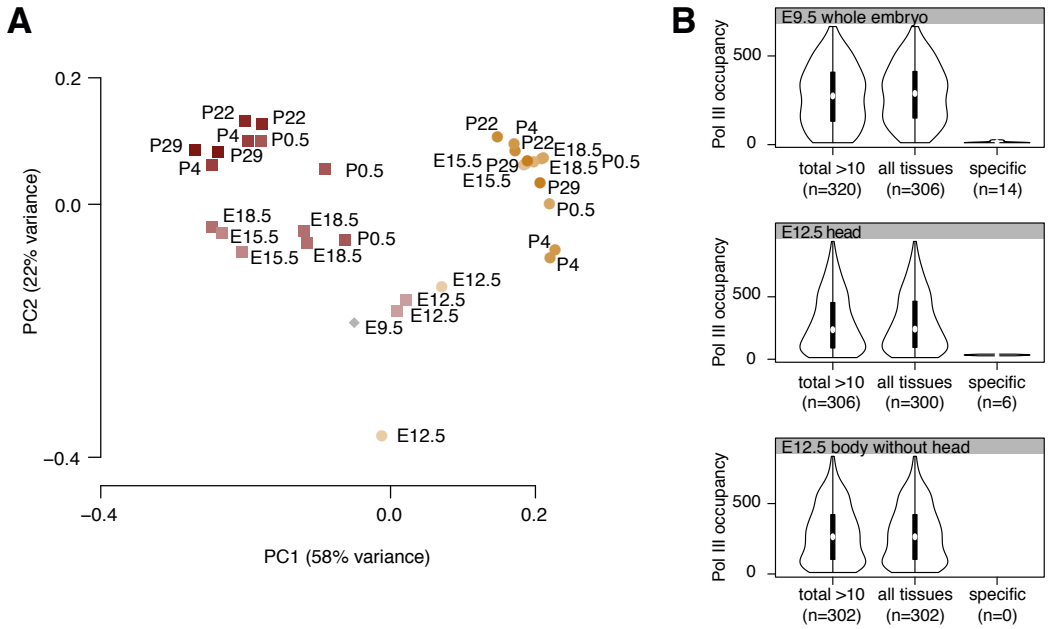

Figure S6. Early developmental stage-specific tRNA genes are lowly expressed.
(**A**) Factorial map of the principal components (PC) analysis of Pol III occupied tRNA gene expression levels in liver (red), brain (yellow), embryonic body without head (light red) and head (light yellow) of stage E12.5, as well as whole E9.5 embryo (gray). The proportion of variance explained by the PC is indicated in parenthesis. This is an updated version of the originally published plot; please refer to https://github.com/klmr/trna/blob/08e99420bf8f1eb92ce1c099cb295eee2b3c2954/README.md.
(**B**) Violin plots represent normalized enrichment of Pol III at tRNA genes identified in E9.5 whole embryo (top), E12.5 head (middle) and E12.5 body without head (bottom) tissue. In parentheses are the numbers of tRNA genes transcribed in the particular embryonic stage ("total >10"), which are subdivided into tRNA genes that can be found in the 12 developmental stages according to Figure 3A ("all tissues") and those that are specific for the embryonic stage ("specific").

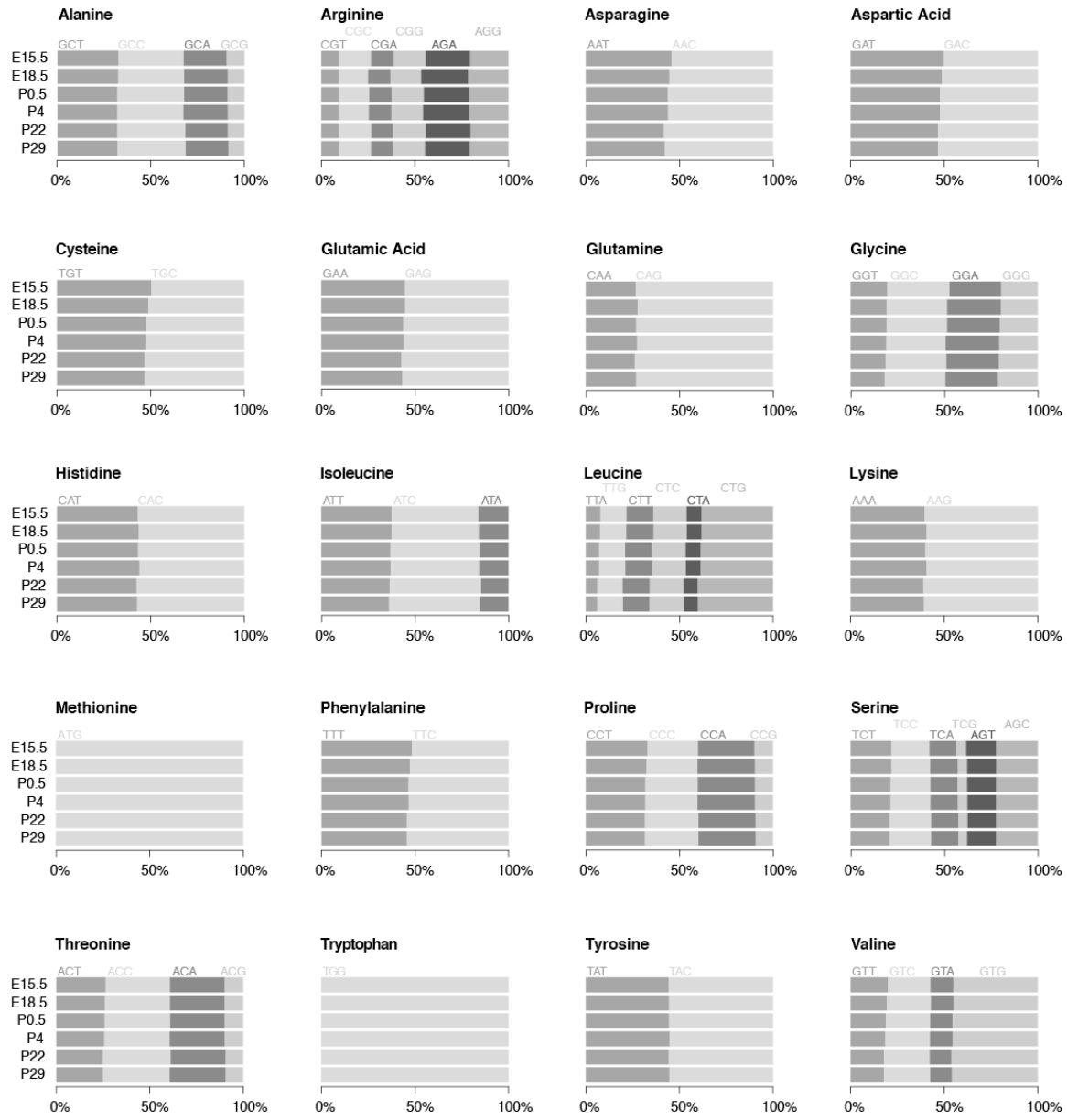

Figure S7: Observed codon usage in mRNA transcriptomes of developing mouse liver
Proportional frequencies weighted by transcript expression are shown for triplet codons ordered by amino acid as a bar plot, where gray shading is by triplet codon. Data are obtained from liver mRNA-seq data of all six developmental stages.

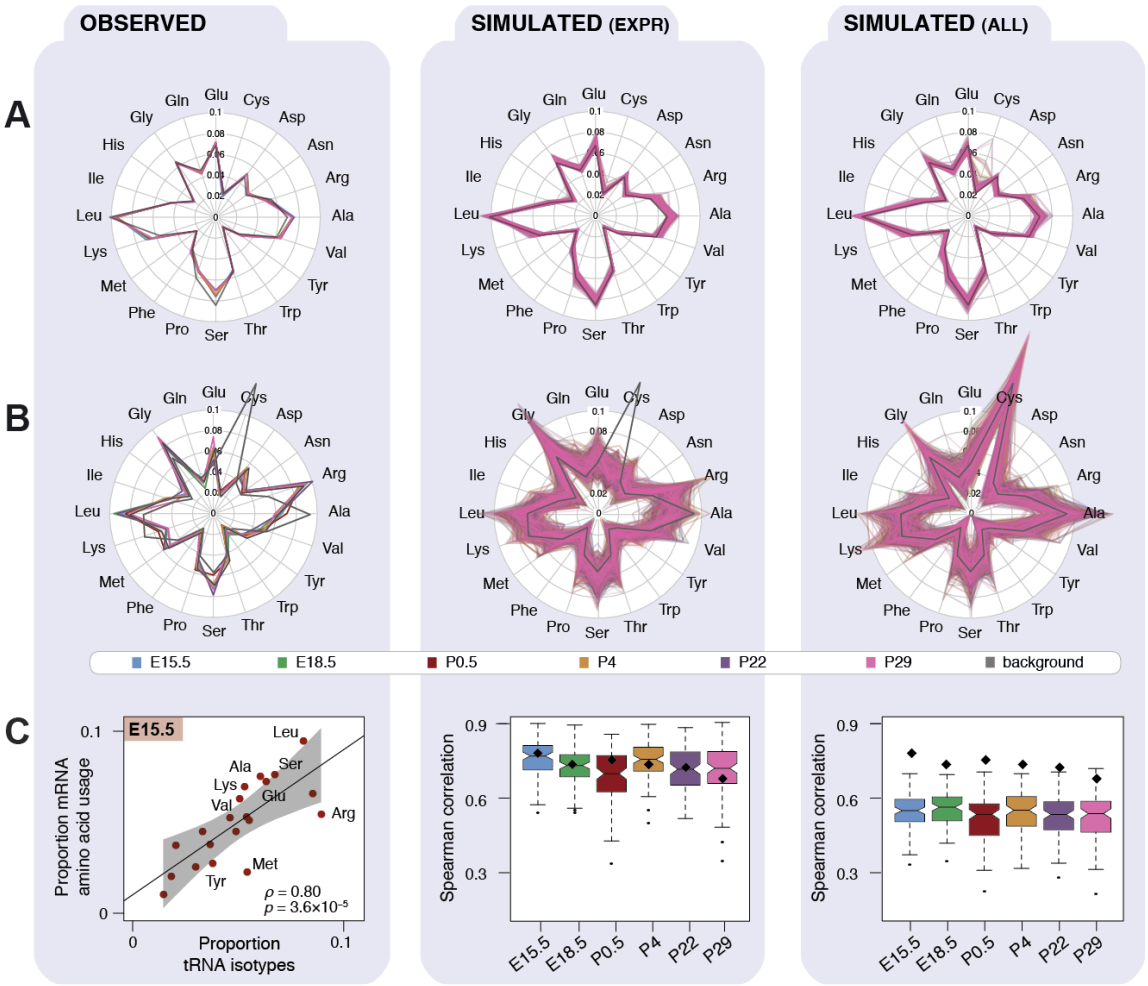

Figure S8: Observed and simulated amino acid and isotype usage in transcriptomes across mouse liver development

Each panel (**A**–**C**) consists of three columns: experimentally observed data (left), simulated patterns of transcription randomized among either the expressed genes (middle) or all genomically encoded genes (right). Transcriptomes of each developmental stage were simulated 100 times (Methods). Proportional frequencies weighted by transcript expression are shown for (**A**) 20 amino acids as a radial plot, where data lines are colored by developmental stage and the background of all genomically annotated mRNA genes is in gray. Labels within grid of radial plot describe ratios. Proportional frequencies weighted by Pol III binding are shown for (**B**) 20 isotypes as a radial plot, both colored as above (gray, background of all genomically annotated tRNA genes). (**C**) Plot right panel shows Spearman's rank correlation coefficients (ρ) and p-values (p) of Pol III binding to tRNA isotypes (x-axis) and transcriptomic amino acid frequencies weighted by expression obtained from mRNA-seq data (y-axis) in E15.5 liver (experimentally observed data) and all six developmental stages (simulated data). Amino acid isotypes outside the 99% confidence interval (gray area within plot in **C** right) are named. Observed Spearman's rank correlation coefficients across all stages (colored as above) are indicated by black diamonds in plot **C** middle and left panels.

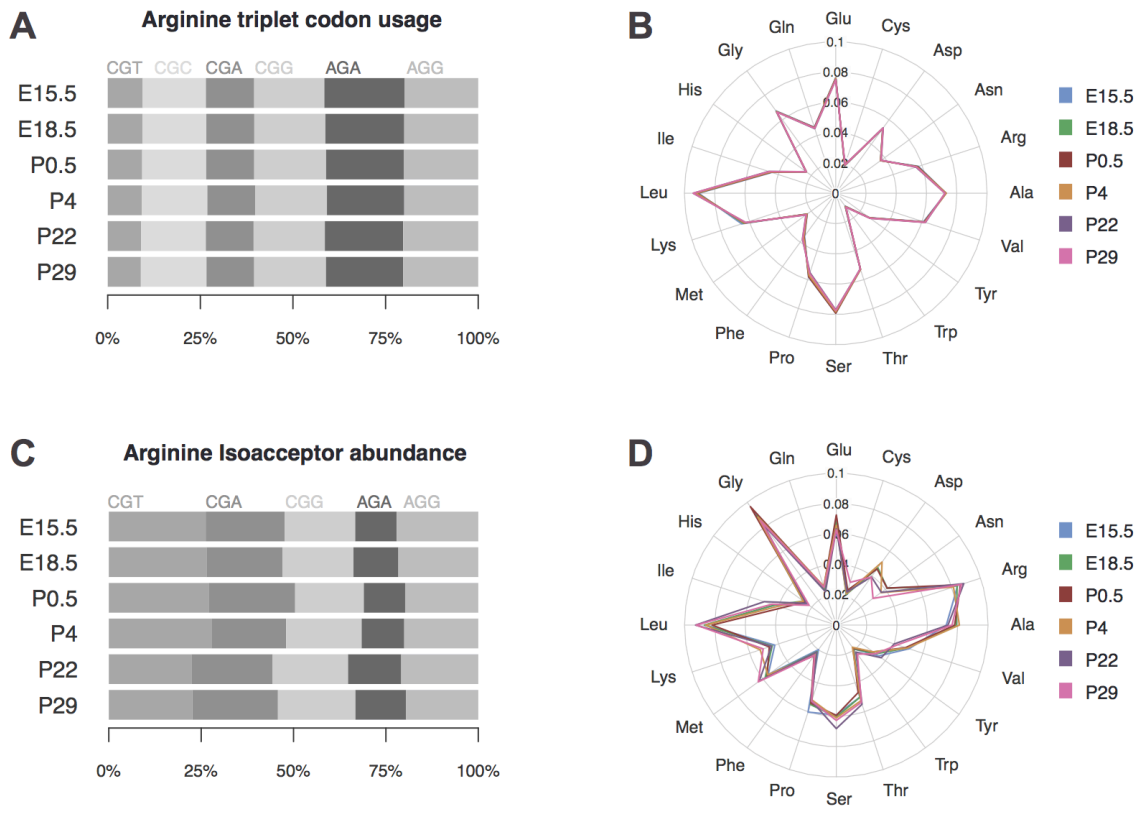

Figure S9. mRNA codon usage and Pol III occupancy of tRNA isotypes in developing mouse brain tissue.
Proportional frequency weighted by transcript expression of (**A**) arginine triplet codons, (**B**) amino acids, (**C**) Pol III binding of arginine isoacceptors and (**D**) Pol III binding of amino acid isotypes. Gray shading in is by triplet codon (**A**) or tRNA anticodon (**C**). Labels within grid of radial plot describe proportions.

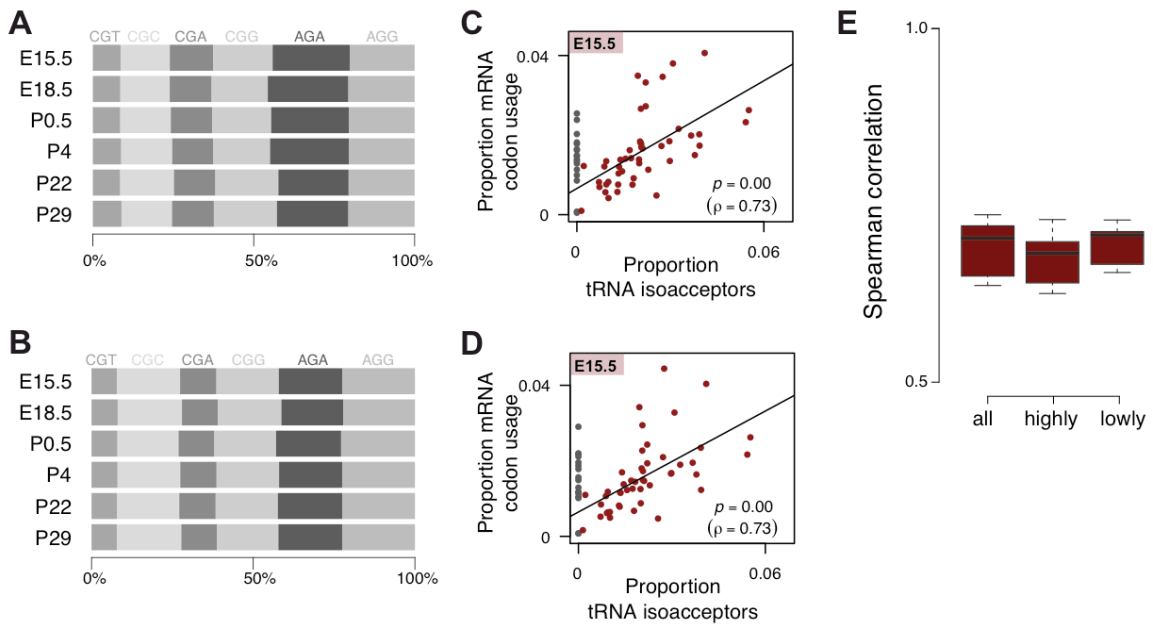

Figure S10. Highly *versus* lowly expressed protein-coding genes show no differential codon usage.

Proportional frequencies weighted by transcript expression are shown for arginine triplet codons as a bar plot of (**A**) highly (90^{th}–95^{th} percentile) and (**B**) lowly expressed (25^{th}–50^{th} percentile) protein-coding genes during liver development, where gray shading is by triplet codon. Plots show Spearman's rank correlation coefficients (ρ) and *p*-values (*p*) of Pol III binding to tRNA isoacceptors (x-axis) and transcriptomic codon frequencies weighted by expression obtained from mRNA-seq data (y-axis) in E15.5 liver of (**C**) highly and (**D**) lowly expressed protein-coding genes. Anticodon isoacceptors (grey dots in plots) are not encoded in the mouse genome and were excluded from calculating the correlation coefficients. (**E**) Variances of correlation values over all stages in liver (i) all expressed protein-coding genes, (ii) highly and (iii) lowly expressed protein-coding gene sets.

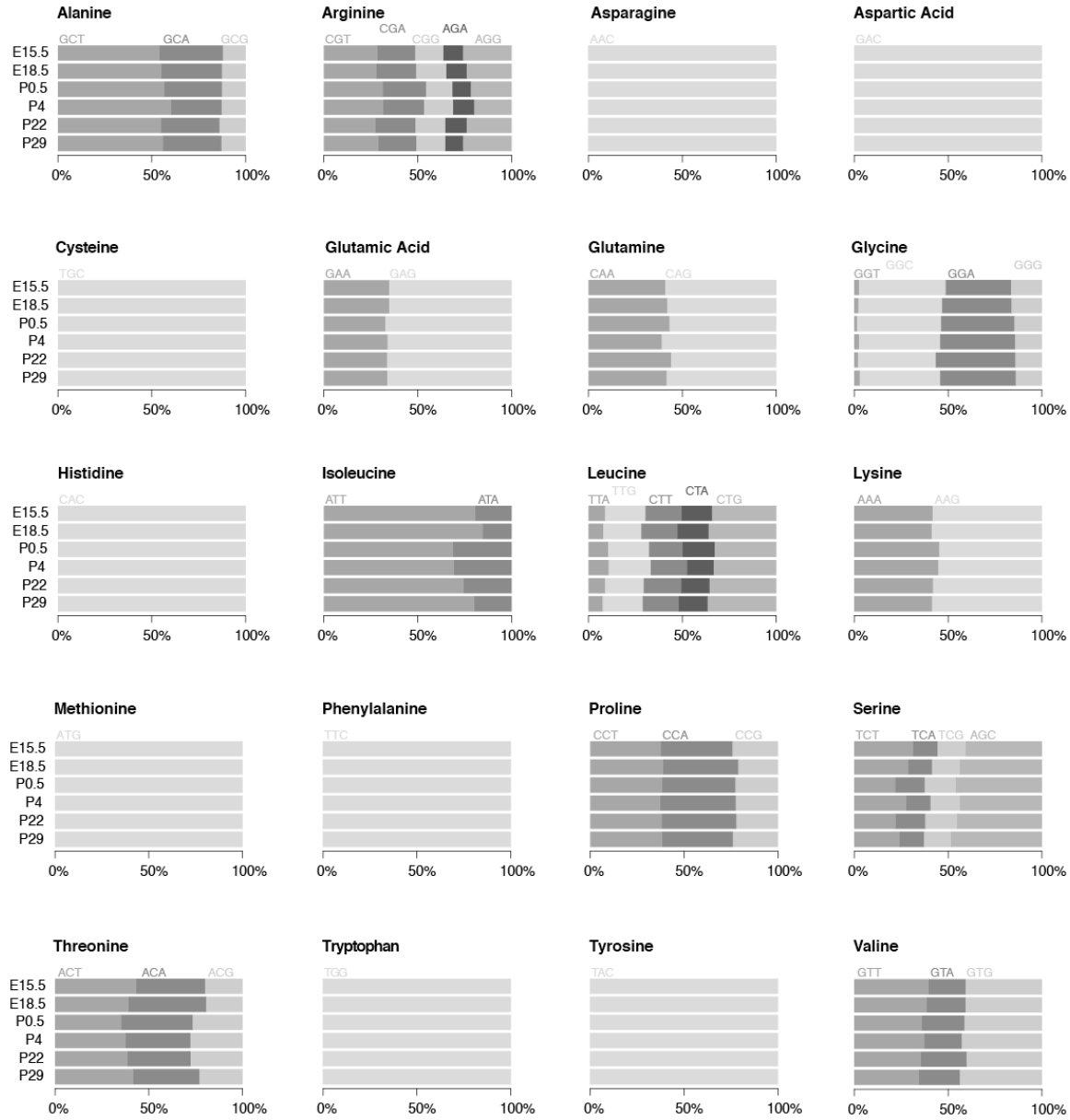

Figure S11: Observed anticodon usage in tRNA transcriptomes of developing mouse liver
Proportional frequencies weighted by by Pol III binding are shown for triplet codons ordered by amino acid as a bar plot, where gray shading is by triplet codon. Data are obtained from liver Pol III ChIP-seq data of all six developmental stages.

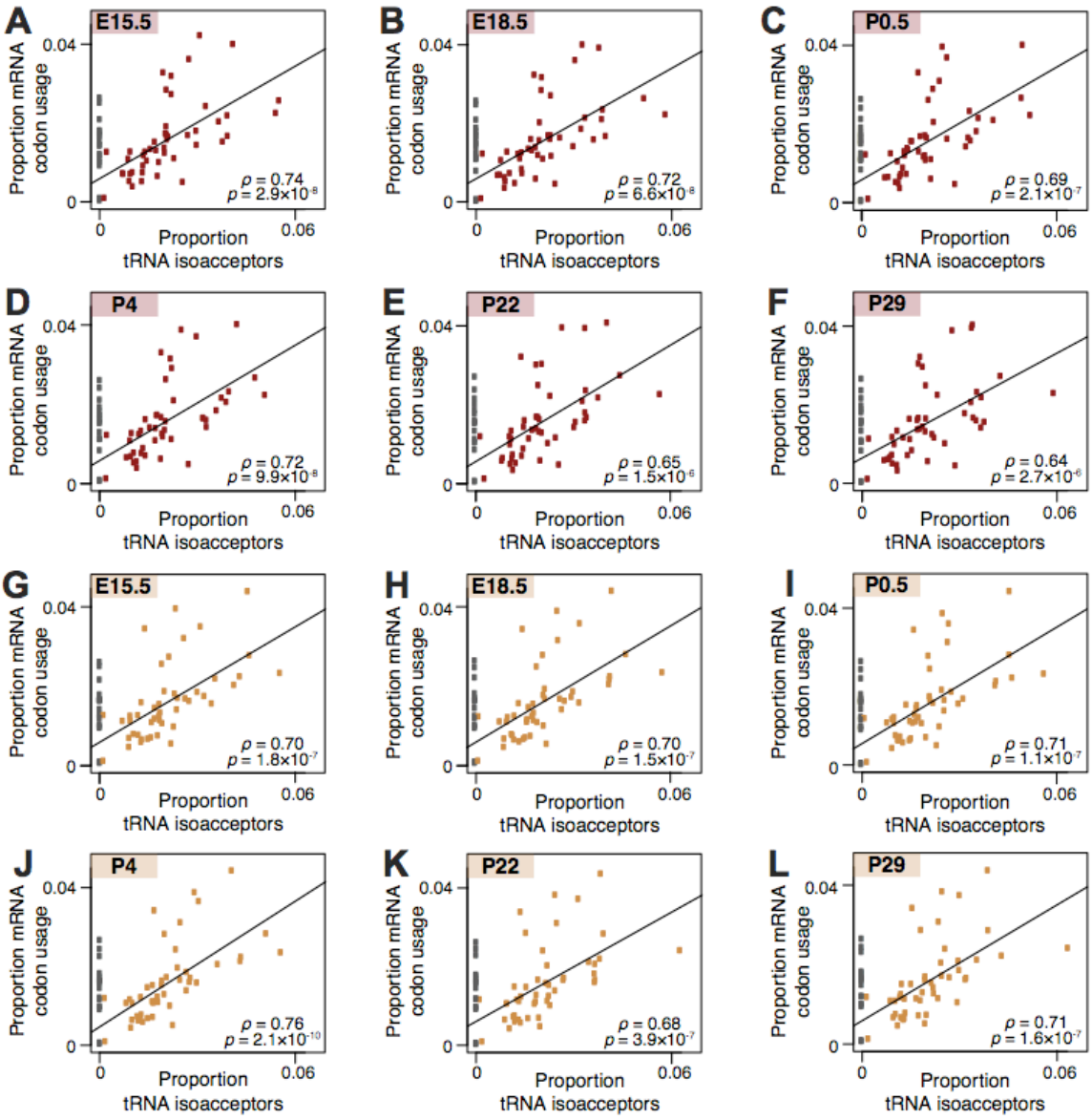

Figure S12. Transcriptomic mRNA codon usage and Pol III binding to tRNA isoacceptors correlate in developing mouse liver and brain.
Plots show correlation of proportional Pol III binding to tRNA isoacceptors (x-axis) and transcriptomic codon frequencies weighted by expression obtained from mRNA-seq data (y-axis). Correlation plots for developing liver (**A**–**F**) and brain (**G**–**L**) are shown. Indexed box in top left indicates developmental stage. Gray dots represent degenerated codons. Spearman's rank correlation coefficients (ρ) are reported along with their *p*-values in bottom right of each panel.

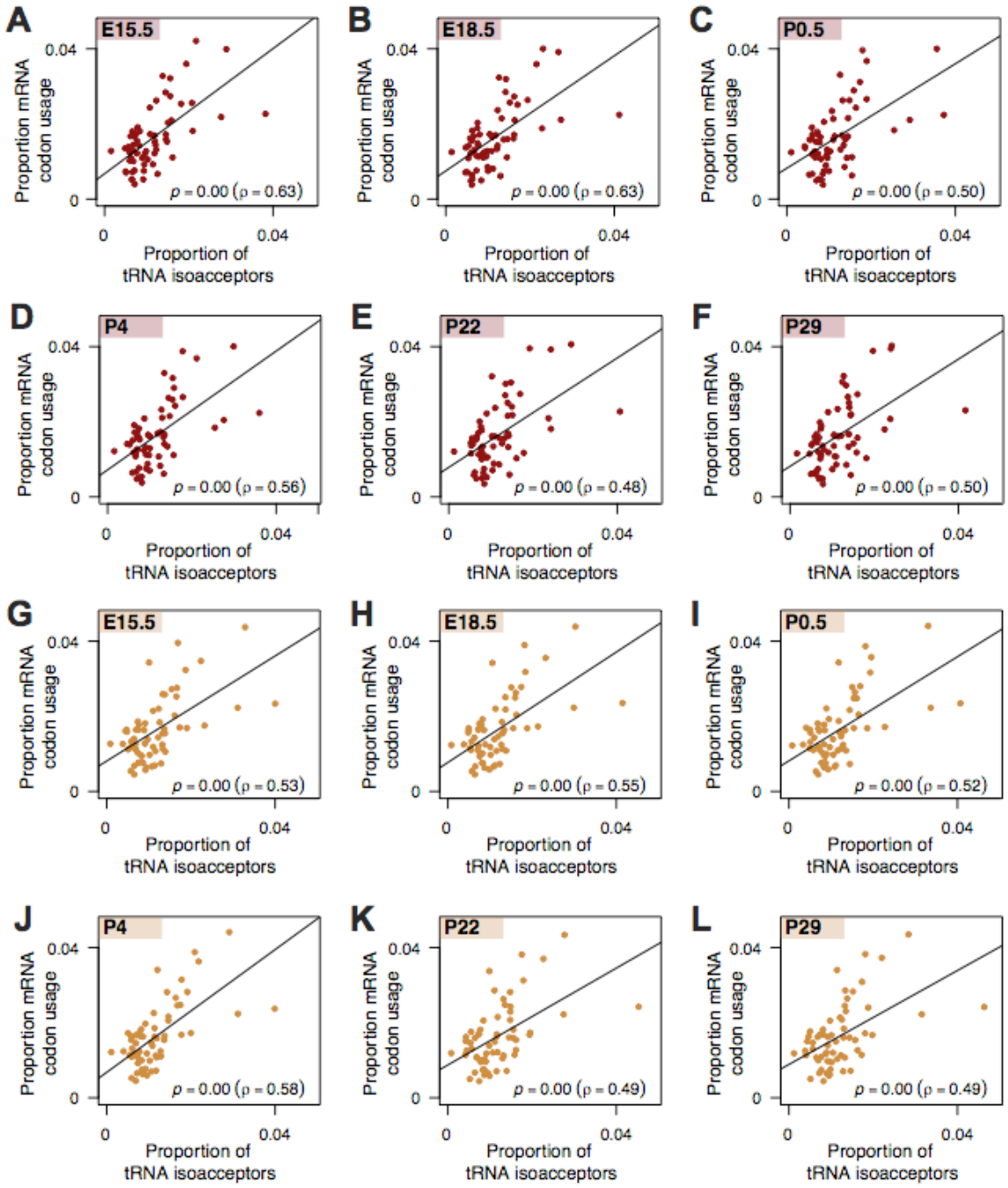

Figure S13. Transcriptomic mRNA codon usage and wobble corrected Pol III binding to tRNA isoacceptors correlate in developing mouse liver and brain.
Plots show correlation of proportional Pol III binding to tRNA isoacceptors corrected according to wobble pairing (x-axis) and transcriptomic codon frequencies weighted by expression obtained from mRNA-seq data (y-axis). Correlation plots for developing liver (**A**–**F**) and brain (**G**–**L**) are shown. Indexed box in top left indicates developmental stage. Spearman's rank correlation coefficients (ρ) are reported along with their *p*-values in bottom right of each panel.

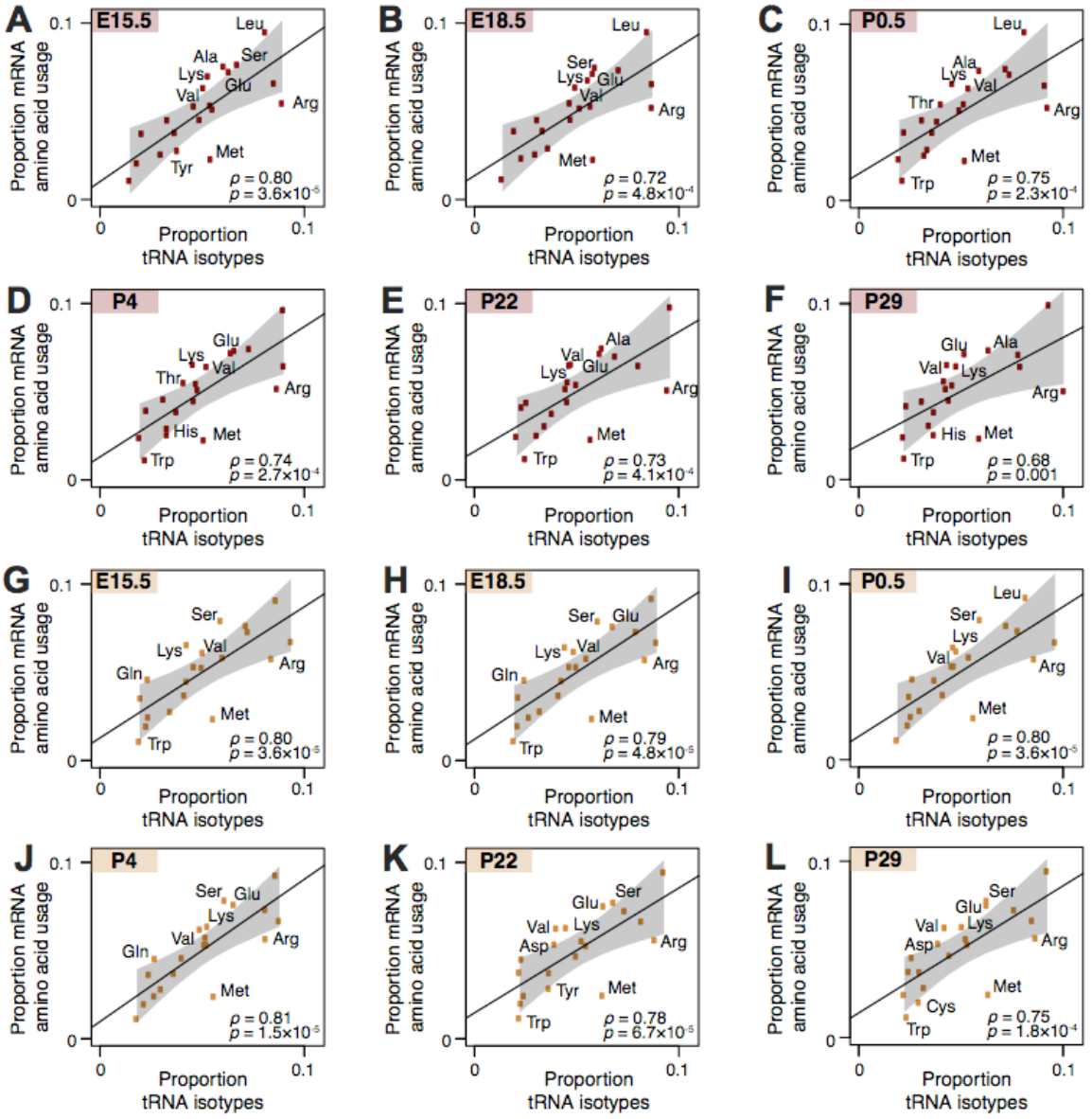

Figure S14. Transcriptomic mRNA amino acid usage and Pol III binding to tRNA isotypes correlate in developing mouse liver and brain.
Plots show correlation of Pol III binding to tRNA isotypes (x-axis) and transcriptomic amino acid frequencies weighted by expression obtained from mRNA-seq data (y-axis). Correlation plots for developing liver (**A**–**F**) and brain (**G**–**L**) are shown. Indexed box in top left indicates developmental stage. Gray area represents 99% confidence interval. Spearman's rank correlation coefficients (ρ) and the corresponding *p*-values (*p*) are reported in top left and bottom right, respectively of each panel. Amino acid isotypes outside the 99% confidence interval (gray area) are named.

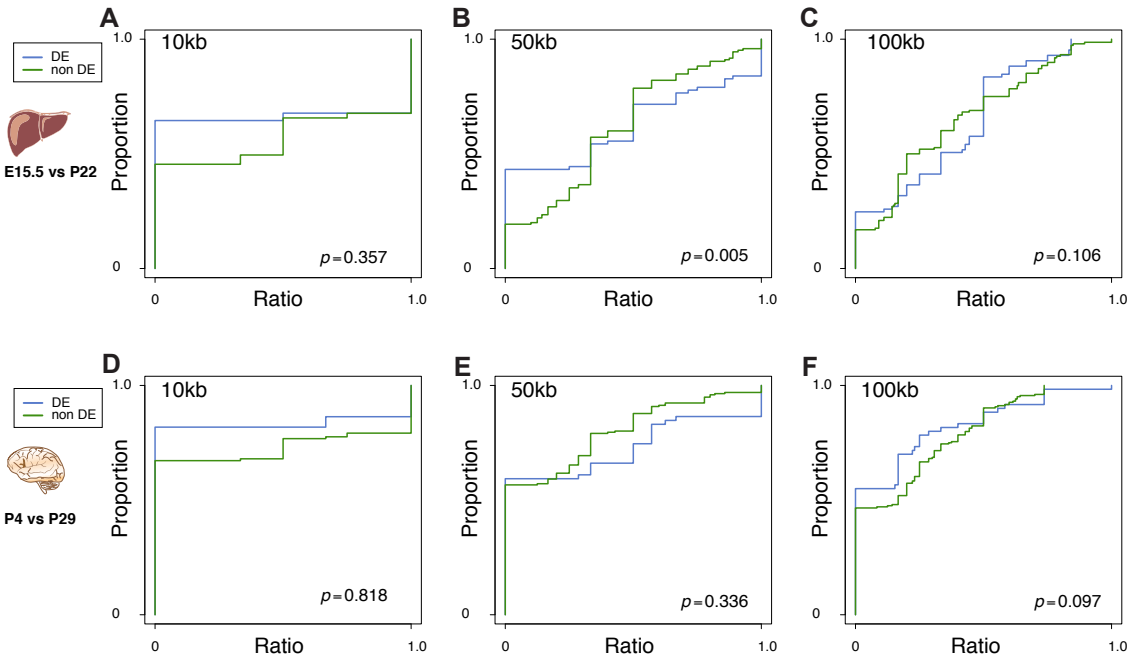

Figure S15. Differentially expressed tRNA genes show no colocalization with differentially expressed protein-coding genes.

In each plot, the blue line is the cumulative distribution of the ratio of the number of upregulated mRNA genes to the number of all mRNA genes, in the neighborhood of each upregulated tRNA gene (Methods). The green line is the cumulative distribution of the ratios of the number of upregulated mRNA genes (FDR cutoff 0.01) to the number of all mRNA genes, in the neighborhood of each tRNA gene that is not differentially expressed. Significant differences between these two distributions reveal situations where upregulated tRNA genes are significantly (by Kolmogorov-Smirnov test) associated with upregulated protein-coding genes. Different window sizes were used, ranging from 10 kb, 50 kb to 100 kb around tRNA genes. Pairwise comparison of (**A–C**) E15.5 and P22 in liver as well as (**D–F**) P4 and P29 in brain are shown. This analysis was repeated using two additional FDR cutoffs (0.05 and 0.1, data not shown). Under the assumption that there was an observable colocalization effect, we would expect there to be a robust signal, i.e. consistent significance across different tested parameters. However, of the eighteen tests, only one was marginally significant (corrected p=0.013), after correcting for multiple testing (Bonferroni), indicating the absence of any strong localization effect.