This document details the steps to generate the data tables that are then provided in Supplemental Tables 2-15. Code for downstream analysis is provided in the R Markdown files **adult_mouse_full_analysis.Rmd** and **fetal_human_full_analysis.Rmd**. All code is run in R or in bash scripts that can be run on a terminal.

Note that several of the raw files are very large (>2 GB), so processing them will require preferably >256 GB of storage space and a significant number of compute hours.

Step 1: Download the tRNA gene annotations for mouse (mm10) and human (hg19)

1. Download the mm10 and hg19 gtRNAdb annotations, which are performed by tRNAscan-SE v2.0
   o hg19: http://gtrnadb.ucsc.edu/genomes/eukaryota/Hsapi19/
   o mm10: http://gtrnadb.ucsc.edu/genomes/eukaryota/Mmusc10/
2. Unzip the download, and extract the files ending in "confidence-set.ss", which will be the ones used to load GRanges in R. These contain the high-confidence tRNA gene annotations for each genome.

Step 2: Download the scRNA-seq datasets for adult mouse and fetal human

- Adult mouse scRNA-seq dataset
  1. Download the Seurat objects from the Tabula Muris Figshare website for the FACS dataset here: https://figshare.com/articles/dataset/Robject_files_for_tissues_processed_by_Seur at/5821263
  2. For each dataset, perform the following. In the Supplemental Code (included in the Supplemental Material) is a script called **sample_scRNA_seq_SCTransform_mRNA_loader.R** that performs these steps is provided, in which the paths of the relevant files must be specified.
     ▪ Extract the count matrix from the Seurat object.
     ▪ Retain only gene entries that are protein-coding. A list of protein-coding genes can be extracted from a GTF in the Ensembl database for mm10: http://useast.ensembl.org/Mus_musculus/Info/Index.
     ▪ Perform SCTransform to variance-stabilize the counts.
  3. Combine the corrected count matrices from all the FACS datasets into a single matrix of cell barcode by gene expression. The cell barcodes are important for linking the cells to their cell type annotation.
  4. Combine the cells together by cell type using the metadata provided by the dataset, which links each cell barcode with its cell type annotation. Note that this is not yet size corrected to account for differences in sequencing depth and number of cells per cell type, which will be done in the processing steps below.
  5. This count matrix can now be used in the processed data script **adult_mouse_full_analysis.Rmd** as the variable **mRNA_gene_expression_per_cell_type** to reproduce the size-corrected data matrix in the supplemental file called **AM_mRNA_gene_expression_per_cell_type_size_corrected.txt**

- Fetal human scRNA-seq dataset
  1. Download the raw gene count sparse matrices and metadata files here: https://descartes.brotmanbaty.org/bbi/human-gene-expression-during-development
  2. For each dataset, perform the following. In the Supplemental Code (included in the Supplemental Material) is a script called **sample_scRNA_seq_SCTransform_mRNA_loader.R** that performs these steps is provided, in which the paths of the relevant files must be specified.
     - Extract the count matrix from the Seurat object.
     - Retain only gene entries that are protein-coding. A list of protein-coding genes can be extracted from a GTF in the Ensembl database for hg19: http://useast.ensembl.org/Homo_sapiens/Info/Index
     - Perform SCTransform to variance-stabilize the counts.
  3. Combine the corrected count matrices from all datasets into a single matrix of cell barcodes by gene expression. The cell barcodes are important for linking the cells to their cell type annotation.
  4. Combine the cells together by cell type using the metadata provided by the dataset, which links each cell barcode with its cell type annotation. Note that this is not yet size corrected to account for differences in sequencing depth and number of cells per cell type, which will be done in the processing steps below.
  5. This count matrix can now be used in the processed data script **fetal_human_full_analysis.Rmd** as the variable **mRNA_gene_expression_per_cell_type** to reproduce the size-corrected data matrix in the supplemental file called **FH_mRNA_gene_expression_per_cell_type_size_corrected.txt**

Step 3: Download the scATAC-seq datasets for adult mouse and fetal human

Note: In these instructions, <string> refers to a filename. For example, <BAM_file> could be foo.bam, etc.

- Adult mouse scATAC-seq dataset
  1. Download the BAM files, count matrix, and metadata files here: https://atlas.gs.washington.edu/mouse-atac/data/.
  2. For each BAM file, perform the following:
     - Convert the BAM files into fragment files using sinto. The following commands were used:
       - sinto fragments -b <BAM_file> -f <fragment_file> --barcode_regex "[^:]*"
     - The fragment files will be in the mm9 coordinates, rather than the desired mm10 coordinates. mm10 is desired for ease of comparison to other datasets, and because of an updated list of tRNA coordinates in the gtRNAdb. Therefore, use the UCSC liftover command to convert mm9 coordinates to mm10 coordinates. The liftover executable is free for personal and non-profit academic research. Note that to perform this

liftover, you must also download the corresponding mm9tomm10 liftover chain [here](). The following commands were used:
- liftOver <fragment file> mm9ToMm10.over.chain <lifted_fragments_file>

- Sort, compress, and index the fragment file as follows:
  - sort -k 1,1 -k2,2n <fragments> > <sorted_fragments>
  - bgzip <sorted_fragments>
  - tabix -p bed <compressed_sorted_fragments>
- A script called **sample_scATAC_seq_tRNA_loader.R** that performs the commands below is provided in the Supplemental Code (within Supplemental Materials). In the script, indicate the paths to the fragment file, the count matrix, and the mm10 gtRNAdb annotation:
  - The fragment file is now ready for use in Signac. Using Signac, create a Fragment object for the fragment file. Create a Chromatin assay in Signac with that Fragment object.
  - Load in the mm10 tRNA Granges (the high confidence tRNA genes in the "confidence-set.ss" file.
  - Use the FeatureMatrix command to identify how many cuts there are to tRNA genes in the Fragment object.
  - Extract the tRNA cuts by cell matrix generated by FeatureMatrix.

3. Combine all tRNA cut per cell matrices together. Use the metadata to keep track of the cell type annotations for each cell using the metadata provided by the dataset.
4. Combine the data belonging to the same cell type together using the metadata provided by the dataset, by linking the cell barcodes to their cell type annotation. Note that this is not yet size corrected to account for differences in sequencing depth and number of cells per cell type, which will be done in the processing steps below.
5. This count matrix can now be used in the processed data script **adult_mouse_full_analysis.Rmd** as the variable **tRNA_gene_expression_per_cell_type** to reproduce the size-corrected data matrix in the supplemental file called **AM_tRNA_gene_expression_per_cell_type_size_corrected.txt**

- Fetal human ATAC-seq dataset
  1. Download the fragment files here: https://descartes.brotmanbaty.org/bbi/human-chromatin-during-development/. Since the files are already formatted as fragment files, no conversion from BAM (as done in the mouse scATAC-seq dataset) is necessary. Additionally, since the genome version is already hg19, no genomic liftovers are necessary either.
  2. For each fragment file, perform the following using the same script **sample_scATAC_seq_tRNA_loader.R** as above for mouse:
     - Using Signac, create a Fragment object for the fragment file. Create a Chromatin assay in Signac with that Fragment object.
     - Load in the hg19 tRNA Granges (the high confidence tRNA genes in the "confidence-set.ss" file.

- Use the FeatureMatrix command to identify how many cuts there are to tRNA genes in the Fragment object.
- Extract the tRNA cuts by cell matrix generated by FeatureMatrix.
3. Combine all tRNA-cut-per-cell matrices together. Use the metadata to keep track of the cell type annotations for each cell using the metadata provided by the dataset.
4. Combine the data belonging to the same cell type together using the metadata provided by the dataset, by linking the cell barcodes to their cell type annotation. Note that this is not yet size corrected to account for differences in sequencing depth and number of cells per cell type, which will be done in the processing steps below.
5. This count matrix can now be used in the processed data script **fetal_human_full_analysis.Rmd** as the variable **tRNA_gene_expression_per_cell_type** to reproduce the size-corrected data matrix in the supplemental file called **FH_tRNA_gene_expression_per_cell_type_size_corrected.txt**