Section 1

Pre-processing pipeline for RNAmappe.R

Data files and folders:

1) Start with a single empty folder called RNAmappe.R recommended to be on the Desktop.
2) Place the RNA-Seq fasta reads for wildtype and mutant in this folder.
3) A reference genome (FASTA) and gene model (GTF) are required.

In the command line, use cd to navigate to the RNAmappe.R directory on your Desktop:

```cd ~/Desktop/RNAmappe.R```

Create environmental variables for the reference genome fasta and gtf files:

Anything with <file> needs to be changed to the name of the actual file located in the directory.

/path/to/<file> represents the absolute path of the file. To retrieve the absolute path of a directory use ```pwd``. Copy the output and then insert <file> at the end of the output.

```refGenome=/path/to/<referenceGenome.Fa>```

```gtfFile=/path/to/<gtfFile.gtf>```

Define mut and wt fasta reads:

```mutRead1=path/to/<mutant_R1.fasta.gz>```

```mutRead2=/path/to/<mutant_R2.fasta.gz>```

```wtRead1=/path/to/<wtR1.fasta.gz>```

```wtRead2=/path/to/<wtR2.fasta.gz>```

Define adapter sequences for mut and wt reads:

mutAdapter1=<adapter_seq>

mutAdapter2=<adapter_seq>

wtAdapter1=<adapter_seq>

wtAdapter2=<adapter_seq>
```

Conda environment and package installation

It is required that conda be installed on the computer being used to perform the RNAmappe.R pipeline. The download and installation information for MacOS can be found at the link below.

https://docs.conda.io/projects/conda/en/latest/user-guide/install/macos.html

A conda environment should be created in the RNAmappe.R folder and should contain all of the package dependencies to run the pipeline.

Required packages:

Package	Version
fastqc	0.12.1
cutadapt	3.5
trimmomatic	0.39
hisat2	2.2.1
samtools	1.18
bcftools	1.17
cufflinks	2.2.1
ensembl-vep	13.0
R	No need to include version

To create this environment: ```conda create --name RNAmappe.R ```

To activate environment: ```conda activate RNAmappe.R```

To install packages: ```conda install <pkg>=<version>```

The environment should be activated before running any part of the pipeline.

To run the preprocessing script:

```bash Trim_Align_Call.sh```
```

****** If aligned .bam files already exist run the Split_chrom.sh script to call variants and split into separate chromosome .vcf files**:

```bash Split_chrom.sh```

Once these files are created, move on to section 2.
```