echo "Creating directory structure..."

mkdir -p ~/rna\_project/{raw\_data,qc\_reports,trimmed\_data,alignment,counts,reference}

echo "Directories created (if they did not exist)

echo "Starting SRA download..."

cd ~/rna\_project/raw\_data

prefetch SRR000111 # Download .sra file

prefetch SRR000112

fastq-dump --split-files SRR000111

fastq-dump --split-files SRR000112

echo "SRA data downloaded and converted to FASTQ."

echo "Running FastQC on raw data..."

fastqc SRR000111\_1.fastq SRR000111\_2.fastq \

SRR000112\_1.fastq SRR000112\_2.fastq \

-o ../qc\_reports

echo "Raw reads QC done. Check the HTML reports in qc\_reports folder."

echo "Trimming reads with Trimmomatic..."

cd ~/rna\_project

trimmomatic PE -threads 4 \

raw\_data/SRR000111\_1.fastq raw\_data/SRR000111\_2.fastq \

trimmed\_data/SRR000111\_1\_paired.fq.gz trimmed\_data/SRR000111\_1\_unpaired.fq.gz \

trimmed\_data/SRR000111\_2\_paired.fq.gz trimmed\_data/SRR000111\_2\_unpaired.fq.gz \

ILLUMINACLIP:$CONDA\_PREFIX/share/trimmomatic/adapters/TruSeq3-PE.fa:2:30:10 \

LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:50

trimmomatic PE -threads 4 \

raw\_data/SRR000112\_1.fastq raw\_data/SRR000112\_2.fastq \

trimmed\_data/SRR000112\_1\_paired.fq.gz trimmed\_data/SRR000112\_1\_unpaired.fq.gz \

trimmed\_data/SRR000112\_2\_paired.fq.gz trimmed\_data/SRR000112\_2\_unpaired.fq.gz \

ILLUMINACLIP:$CONDA\_PREFIX/share/trimmomatic/adapters/TruSeq3-PE.fa:2:30:10 \

LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:5

echo "Trimming completed. Paired reads are in trimmed\_data/."

echo "Downloading and decompressing reference genome..."

cd ~/rna\_project/reference

wget ftp://ftp.ensembl.org/pub/release-109/fasta/homo\_sapiens/dna/Homo\_sapiens.GRCh38.dna.primary\_assembly.fa.gz

gunzip Homo\_sapiens.GRCh38.dna.primary\_assembly.fa.gz

echo "Reference genome downloaded and unzipped."

echo "Building HISAT2 index..."

hisat2-build Homo\_sapiens.GRCh38.dna.primary\_assembly.fa GRCh38\_index

echo "HISAT2 index built. Files: GRCh38\_index.\*"

echo "Aligning reads..."

cd ~/rna\_project

mkdir -p alignment

hisat2 -p 4 -x reference/GRCh38\_index \

-1 trimmed\_data/SRR000111\_1\_paired.fq.gz \

-2 trimmed\_data/SRR000111\_2\_paired.fq.gz \

-S alignment/SRR000111.sam

hisat2 -p 4 -x reference/GRCh38\_index \

-1 trimmed\_data/SRR000112\_1\_paired.fq.gz \

-2 trimmed\_data/SRR000112\_2\_paired.fq.gz \

-S alignment/SRR000112.sam

echo "Alignment done. SAM files are in alignmen

echo "Converting, sorting, and indexing alignments..."

cd alignment

samtools view -bS SRR000111.sam | samtools sort -o SRR000111.sorted.bam

samtools index SRR000111.sorted.bam

samtools view -bS SRR000112.sam | samtools sort -o SRR000112.sorted.bam

samtools index SRR000112.sorted.bam

echo "Sorted BAM files ready. Index files created."

echo "Downloading annotation (GTF) for featureCounts..."

cd ~/rna\_project/reference

wget ftp://ftp.ensembl.org/pub/release-109/gtf/homo\_sapiens/Homo\_sapiens.GRCh38.109.gtf.gz

gunzip Homo\_sapiens.GRCh38.109.gtf.gz

cd ~/rna\_project

mkdir -p counts

echo "Running featureCounts..."

featureCounts -T 4 \

-a reference/Homo\_sapiens.GRCh38.109.gtf \

-o counts/all\_samples\_counts.txt \

alignment/SRR000111.sorted.bam \

alignment/SRR000112.sorted.bam

echo "featureCounts completed. Output saved to counts/all\_samples\_counts.txt."

echo "RNA-seq pipeline completed successfully!"