Guide to Run the Shell Script

To run the shell script, follow these steps:

## Step 1: Prepare the Script

1. Create a new shell script file:  
 - Open a terminal and navigate to the directory where you want to save the script.  
 - Use a text editor (e.g., `nano`, `vim`, or any GUI-based editor) to create a new file, for example:

nano run\_pipeline.sh

2. Copy and paste the script:  
 - Copy the shell script I provided in the previous message and paste it into the file.  
 - Save the file and exit the editor (in `nano`, press `CTRL + X`, then `Y` to confirm, and `Enter` to save).

3. Make the script executable:  
 - Run the following command to make the script executable:

chmod +x run\_pipeline.sh

## Step 2: Execute the Script

1. Navigate to the working directory:  
 - Use `cd` to navigate to the directory where your FASTQ files and the script are located:

cd /path/to/your/working/directory

2. Run the script:  
 - Execute the script by typing the following command in the terminal:

./run\_pipeline.sh

## Important Notes

1. File Availability:  
 - Ensure that all required files (e.g., `.fastq.gz`, `reference\_genome.fa`, `annotation.gtf`) are present in the directory or use appropriate file paths.

2. Script Output:  
 - The script will:  
 - Generate MD5 checksums and store them in `md5\_checksums.txt`.  
 - Unzip the `.fastq.gz` files.  
 - Run FastQC and store results in the `QC\_reports` directory.  
 - Generate a MultiQC report.  
 - Trim the reads using Fastp.  
 - Align the reads with Hisat2 and output SAM files.  
 - Convert SAM to BAM, sort BAM files, and generate alignment statistics.  
 - Run featureCounts for gene quantification.

3. Monitor the Process:  
 - You can monitor the script execution and check the output files in real-time. If needed, add `echo` statements in the script to provide more feedback during execution.