

Exercise: Primer Removal with Cutadapt in Docker

Step 1: Create a new directory for your bioinformatics project:

```
mkdir primer-removal-docker
cd primer-removal-docker
```

Step 2: Install Cutadapt Script

Create a simple script named `remove_primers.sh`:

```
#This script uses Cutadapt to remove primers from the input
sequences.
#!/bin/bash
cutadapt -a ATCGATCG -o sequences_without_primers.fasta
sequences.fasta
```

Step 3: Create Dockerfile

Create a Dockerfile to build the Docker image:

```
# Dockerfile
FROM ubuntu:23.04

LABEL maintainer="george githinji <ggithinji@kemri-wellcome.org>"

# Install necessary dependencies
RUN apt-get update && apt-get -y upgrade && \
    apt-get install -y python3-pip python3-dev build-essential

# install cut-adapt in /usr/local/bin
RUN apt install -y cutadapt

# Cleanup unnecessary packages and files
RUN apt-get remove -y build-essential wget && \
    apt-get autoremove -y && \
    apt-get clean

# Copy the primer removal script into the container
COPY remove_primers.sh /usr/local/bin/remove_primers

RUN chmod +x /usr/local/bin/remove_primers

# Set the working directory
WORKDIR /data

# Entry point command
ENTRYPOINT ["bash"]
```

Step 4: Build the Docker Image

Build the Docker image using the following command:

```
docker build -t primer-removal .
```

Step 5: Run the Docker Container

Run the Docker container, mounting the local directory containing the input sequences into the container:

```
docker run -v $(pwd):/data primer-removal
```

This command mounts the current directory into the `/data` directory in the container and runs the primer removal script.

Step 6: Verify Output

Check the output file (`sequences_without_primers.fasta`) to ensure that primers have been successfully removed.