

FastQC Reporter User Manual

Introduction

Fastqc reporter is a Command Line Interface (CLI) tool built to parse fastqc files into the available sections and generate reports containing the section contents. It also generates a graphical representation of the section and a flag file containing the result of the QC test, which can be the values **pass**, **fail**, or **warn**.

Installation

FastQC Reporter runs in a Python environment (version 3.9 or later) and requires a few dependencies. We will use the Conda environment manager for simplicity:

1. **Install Conda:** Visit <https://docs.conda.io/en/latest/miniconda.html> to download and install Miniconda (a lightweight version of Conda).
2. **Create and Activate Environment:** Open a terminal (on Windows, you can use the "Anaconda Prompt") and create a new environment by typing the following command:

```
>> conda create -c conda-forge -n fastqc_env seaborn pandas matplotlib
```

This command creates a virtual environment named "fastqc_env" and installs the required dependencies.
3. **Activate Environment:** Once the environment is created, activate it by typing:

```
source activate fastqc_env
```

 # Use "activate fastqc_env" for Windows

Running FastQC Reporter

FastQC Reporter is designed to be used directly from the command line. Below, we explain how to do this.

1. **Command Syntax:**

```
python3 fastqc_reporter.py [path_to_fastqc_file] [output_folder] [options]
```

- o **[path_to_fastqc_file]:** Path to your FastQC result file (e.g., `./data/fastqc_data1.txt`).
- o **[output_folder]:** Path where you want to save the generated reports and plots.
- o **[options]:** Optional parameters that specify which sections of the FastQC report to process. The tool can generate plots and reports for specific sections of the FastQC data and for all sections.

2. **Example Usage:** Suppose you have a FastQC output file named `fastqc_data1.txt` in a folder called "data," and you want the results saved in a folder named "solution":

3.

```
python3 fastqc_reporter.py ./data/fastqc_data1.txt ./solution/ -a
```

This command generates reports for the entire FastQC file and saves them in the specified output folder.

Command-Line Options

For a full list of options, run:

```
python3 fastqc_reporter.py -h
```

This will display all available options and their descriptions.

Expected Output

When you run FastQC Reporter, you can expect the following outputs in your specified output folder:

1. **Text Reports** (`report.txt`): Contain detailed information for each section (e.g., quality scores, sequence duplication levels).
2. **Flag Files** (`flag.txt`): Indicate the quality status for each section (values can be "pass," "warn," or "fail").
3. **Plots** (`plot.png`): Visual graphs for each data section (e.g., quality score distributions, GC content).

Tips for Non-Technical Users

- **Terminal Use:** The terminal or command prompt is where you type commands. If you are on Windows, using the "Anaconda Prompt" will make it easier to work with Conda.
- **Directory Navigation:** You may need to change directories (folders) to where your FastQC file is located. For example:
`cd path/to/your/folder`
- **Help Option:** If you need guidance on the available options, the `-h` flag provides a helpful summary.