

Short Read Sequencing Analysis Workshop

Day 4 – FastQC and Trimmomatic

FastQC

This program is used to visualize the quality of the sequencing data before further analysis is completed

The module to use on vieques is: fastqc_0.11.2

The output will be a .html file that will need to be transferred to your local machine to view

Usage statement

```
fastqc -o <output directory> <input file>
```

Trimmomatic

This program is used to trim poor quality bases from the reads

The module is a java -jar program, and you must know the path

On vieques it is /opt/trimmomatic/0.32/trimmomatic-0.32.jar

The Illumina adapters that are pre-programmed into Trimmomatic can be found in the

/opt/trimmomatic/0.32/adapters directory

This is the path name to use for the ILLUMINACLIP option

Single end usage statement

```
java -jar <path_to_trimmomatic> [SE | PE] [-phred33 | -phred64]  
<input_R1.fastq> <R1_output.fastq>  
ILLUMINACLIP:<path_to_trimmomatic_adapters>:<seedmismatches>:<palindrom clip threshold>:<simple clipthreshold>
```

Paired end usage statement

```
java -jar <path_to_trimmomatic> [SE | PE] [-phred33 | -phred64]  
<input_R1.fastq> <input_R2.fastq> <R1_paired_output.fastq>  
<R1_unpaired_output.fastq> <R2_paired_output.fastq>  
<R2_unpaired_output.fastq> ILLUMINACLIP:<path_to_trimmomatic_adapters>  
>:<seedmismatches>:<palindrom clip threshold>:<simple clipthreshold>
```

Options

ILLUMINACLIP	Cut adapter and other illumina-specific sequences from the read.
LEADING	Cut bases off the start of a read, if below a threshold quality
TRAILING	Cut bases off the end of a read, if below a threshold quality
SLIDINGWINDOW	Performs a sliding window trimming approach. It starts scanning at the 5" end and clips the read once the average quality within the window falls below a threshold
MINLEN	Drop the read if it is below a specified length