|  |
| --- |
| ***Insert Laboratory Specific Name Here*** |
| **Read Trimming of Next-Generation Sequencing Data Training Form** |

|  |  |
| --- | --- |
| **Employee Name** | **Training Start Date** |
|  |  |

**Section I – *TELL* - Base Knowledge (Video and Reading Requirements)** *[select videos and documents relevant to your lab processes; add other videos and documents as appropriate]*

|  |  |  |
| --- | --- | --- |
| **Document Name** | **Trainee Initials** | **Date Read** |
| [***An Extensive Evaluation of Read Trimming Effects on Illumina NGS Data Analysis***](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0085024) |  |  |
| [***Read Trimming: Introduction – Phred Scores***](https://drive5.com/usearch/manual/quality_score.html) |  |  |
| [***Read QA and Cleaning***](https://bioinf.comav.upv.es/courses/sequence_analysis/read_cleaning.html) |  |  |
| [***Sickle : Quality trimming Illumina paired-end reads***](https://metagenomics-workshop.readthedocs.io/en/latest/reads-qc/qtrim.html) |  |  |
| [***Trimmomatic : Trimming and Filtering***](https://datacarpentry.org/wrangling-genomics/03-trimming/index.html)  |  |  |
| [***CutAdapt : User Guide***](https://cutadapt.readthedocs.io/en/stable/guide.html) |  |  |
| [***BBDuk : Usage Guide***](https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/usage-guide/) |  |  |

|  |  |  |
| --- | --- | --- |
| **Video Name** | **Trainee Initials** | **Date Watched** |
| [***Trimming***](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0085024) ***Adapters from Fastq Reads*** |  |  |

**Section II – *SHOW* - Observation: Trainee observes the trainer perform all steps in the read trimming SOP**

|  |  |  |
| --- | --- | --- |
| **Discussion Points** | **Trainer Initials** | **Date** |
| What is a PHRED or Q score? |  |  |
| How can we assess the quality of reads before determining whether to trim? |  |  |
| How can quality trimming improve average sequence quality? (What is the net effect of trimming) |  |  |
| *Add additional questions the trainer should ask the trainee to determine level of understanding specific to your protocol.* |  |  |

**Section III – *DO* - Performance under Supervision: Trainee performs all steps in the read trimming SOP under direct trainer supervision**

Controls and/or sample(s) will be provided to the trainee. The trainee will:

*(Find example steps below, insert steps specific to your lab)*

1. *Open a Sequence File in Trimming tool*
2. *Appropriately toggle required parameters with associated values*
3. *Properly saved the trimmed reads in the appropriate archive*

Successful performance criteria: All steps to perform read trimming are followed appropriately, data are accurately processed and saved correctly.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Performance Assessment** | **Yes** | **No** | **Trainer Initials** | **Date** |
| *Sequencing file opened in Trimming tool*  | o | o |  |  |
| *Appropriate parameters and values are applied* | o | o |
| *Resulting data is properly saved following directory structure* | o | o |
| **Comments:** |

**Section IV – *APPLY* - Independent Performance: Trainee individually executes all steps in the read trimming SOP**

Sample(s) will be provided to the trainee. The trainee will:

*(Find example steps below, insert steps specific to your lab)*

1. *Open a Sequence File in Trimming tool*
2. *Appropriately toggle required parameters with associated values*
3. *Properly saved the trimmed reads in the appropriate archive*

Successful performance criteria: All steps to perform read trimming are followed appropriately, data are accurately processed and saved correctly.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Performance Assessment** | **Yes** | **No** | **Trainer Initials** | **Date** |
| *Sequencing file opened in Trimming tool*  | o | o |  |  |
| *Appropriate parameters and values are applied* | o | o |
| *Resulting data is properly saved following directory structure* | o | o |
| **Comments:** |

**Section V – Employee Attestation**

|  |  |  |  |
| --- | --- | --- | --- |
| **Attestations** | **Yes** | **No** | **Trainee Initials** |
| I read and understand the procedures listed in the required reading. | **o** | **o** |  |
| I had an opportunity to discuss my questions with the trainer. | **o** | **o** |  |
| I am satisfied with the explanations provided to me; all my questions were answered. | **o** | **o** |  |

**Section VI – Review and Signatures**

|  |  |  |
| --- | --- | --- |
| **Trainee Name** | **Signature** | **Date** |
|  |  |  |
| **Trainer Name** | **Signature** | **Date** |
|  |  |  |
| **CLIA Technical or General Supervisor (as applicable)** | **Signature** | **Date** |
|  |  |  |

**Appendix A – Trainer Discussion Topic Answer Sheet**

**What is a PHRED or Q score?**

The quality score of a given nucleotide base, given as an integer value, and representing the estimated probability that the base was called incorrectly.

**How can we assess the quality of reads before determining whether to trim?**

Length distribution, PHRED quality distribution, nucleotide frequencies and complexity can all be indicators of read quality. Scores such as mean quality, Q30 (the percentage of bases from the reads that have a PHRED score equal to or greater than 30) among others can provide an idea of the overall sequence quality.

**How can quality trimming improve average sequence quality? (What is the net effect of trimming)?**

Read trimming aims to remove low quality bases by removing only low quality regions, thus resulting in an improved average quality. The fundamental principle of read trimming is to operate an educated estimate of read error rates trying to keep the longest possible high quality subsequence.

***\*Insert answers to additional discussion topics here (including tool-specific questions) \****