|  |
| --- |
| ***Insert Laboratory Specific Name Here*** |
| **Short Read Alignment 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; see examples of tool-specific line items in blue below]*

|  |  |  |
| --- | --- | --- |
| **Document Name** | **Trainee Initials** | **Date Watched** |
| [***Short read alignment: an introduction***](https://www.sevenbridges.com/short-read-alignment-algorithms/) |  |  |
| [***Practical 2 : Short Read Alignment***](https://bioinformatics-core-shared-training.github.io/cruk-summer-school-2018/Introduction/SS_DB/Materials/Practicals/Practical2_alignment_SS.html) |  |  |
| [***BWA Tutorial***](https://icb.med.cornell.edu/wiki/index.php/Elementolab/BWA_tutorial) |  |  |
| [***Samtools Tutorial***](http://quinlanlab.org/tutorials/samtools/samtools.html) |  |  |
| [***Bowtie2***](http://bowtie-bio.sourceforge.net/bowtie2/manual.shtml) |  |  |

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

|  |  |  |
| --- | --- | --- |
| **Discussion Points** | **Trainer Initials** | **Date** |
| What factors can make short read alignment difficult and/or computationally intensive? What methods can help address these challenges? |  |  |
| What does the indexing step achieve with regards to the overall alignment process? |  |  |
| What is the purpose of performing an alignment within the context of the provided workflow? |  |  |
| ***For which type of file does SAMStat provide statistics and QC?*** |  |  |
| *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 short read alignment 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. *Acquire/select sample sequence file(s) for alignment*
2. *Acquire/select/prepare reference genome(s)*
3. *Perform short read alignment with tool of your choosing (e.g. BWA, Bowtie)*
4. *Properly save results in appropriate directories*

Successful performance criteria: All steps to perform short read alignment run are followed appropriately, reports are accurately interpreted and saved correctly.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Performance Assessment** | **Yes** | **No** | **Trainer Initials** | **Date** |
| *Acquired and selected appropriate sample file(s) for alignment* | o | o |  |  |
| *Acquired, selected and prepared appropriate reference genome(s)* | o | o |
| *Executed short read alignment steps as specified in the SOP* | o | o |
| *Saved Results to appropriate directories* | o | o |
| **Comments:** | | | | |

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

Previously run sample(s) will be provided to the trainee. The trainee will:

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

1. *Acquire/select sample sequence file(s) for alignment*
2. *Acquire/select/prepare reference genome(s)*
3. *Perform short read alignment with tool of your choosing (e.g., BWA, Bowtie)*
4. *Properly save results in appropriate directories*

All steps to perform short read alignment run are followed appropriately, reports are accurately interpreted and saved correctly.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Performance Assessment** | **Yes** | **No** | **Trainer Initials** | **Date** |
| *Acquired and selected appropriate sample file(s) for alignment* | o | o |  |  |
| *Acquired, selected and prepared appropriate reference genome(s)* | o | o |
| *Executed short read alignment steps as specified in the SOP* | o | o |
| *Saved Results to appropriate directories* | 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 factors make alignment difficult and/or computationally intensive? What methods can help address these challenges?**

Short read alignment can be difficult when the reference genomes utilized are large, as this can be more computationally demanding and require more time. Additionally, because the matches being performed aren’t exact, the amount of possible candidate matches can be very large and difficult for brute force methods to handle. Improved searching, sequence compression and dynamic programming can all help improve the speed of assembly processes.

**What does the indexing step achieve with regards to the overall alignment process?**

Indexing allows aligners to narrow down the potential origin of a query sequence within the genome. Once an index is created, the original FASTA files are no longer needed by the aligner (e.g., Bowtie2) to align reads to that reference.

**What is the purpose of performing an alignment within the context of the provided workflow?**

*\*The answer to this question will depend on your lab’s workflow, and should account for downstream analyses that will utilize the assembly and overall focus of the pipeline or workflow\**

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

***For which type of file does SAMStat provide statistics and QC?***

*SAMStat is used to assess the quality of a SAM or BAM file?*