STRaM Analysis Workflow

Ver. 14

Wei Guo, 4/11/2022

1. Import STRaM analysis workflow v9 into your galaxy account

|  |
| --- |
|  |
|  |
| Note: if you imported the STRaM analysis workflow before, please skip the step.  Current workflow version: **STRaM\_analysis\_workflow\_v18.ga**  Location: Lab\_shared\_restricted\Bioinformatics\reference genome\STR\workflow |

1. Upload (or copy) all the datasets into a NEW history in Galaxy Server. You may have to use both choices below.

|  |
| --- |
| Choice 1: upload datasets from your local drives |
|  |
|  |
| Choice 2: copy datasets from histories of your galaxy account |
|  |
|  |
|  |
| IMPORTANT NOTE: the order of datasets uploaded or copied/pasted has to be the same as that shown in the right panel above or below. Otherwise, an error message will show up during the workflow run.   1. Paired-end read 1 2. Paired-end read 2 3. Reference human genome hg38 (**GRCh38\_no\_alt\_analysis\_set**) 4. 5’ reference flanking sequences of markers (**STRaM\_v10XY\_flank5\_v3.fasta**) 5. 3’ reference flanking sequences of markers (**STRaM\_v10XY\_flank3\_v3.fasta**) 6. A tabular of markers’ name and genomic locations (**STRaM\_markers\_v10\_list\_v16.tabular**) |

1. Run STRaM analysis workflow.

|  |
| --- |
|  |
|  |

1. Select the uploaded datasets from step 2 into corresponding input items.

|  |
| --- |
| Please ensure the 6 datasets are selected correctly. Then click Run workflow. |
|  |

1. Monitor the progress of the workflow

|  |
| --- |
|  |
|  |

1. Collect results from STR analysis, Flank analysis and XY/mutation analysis.

|  |
| --- |
|  |
| **Workflow outputs**:   1. Comparison between STR analysis and flanking analysis is located in **item 43**; 2. XY and mutation analysis results are located in **item 32.**   Interpretation of tabular data:  “**True**” in column 13 indicates consistent STR lengths between STR and flank analysis;  “**False**” in column 13 indicates different STR lengths between STR and flank analysis; |
|  |

1. Manually correct inconsistent results of the individual markers.

|  |
| --- |
|  |
| The program “**select**” is located under the section of **filter and sort**.  You may use the genomic location of STR start or flank end sequences as searching pattern to pick out the inconsistent STR analysis results. |
|  |