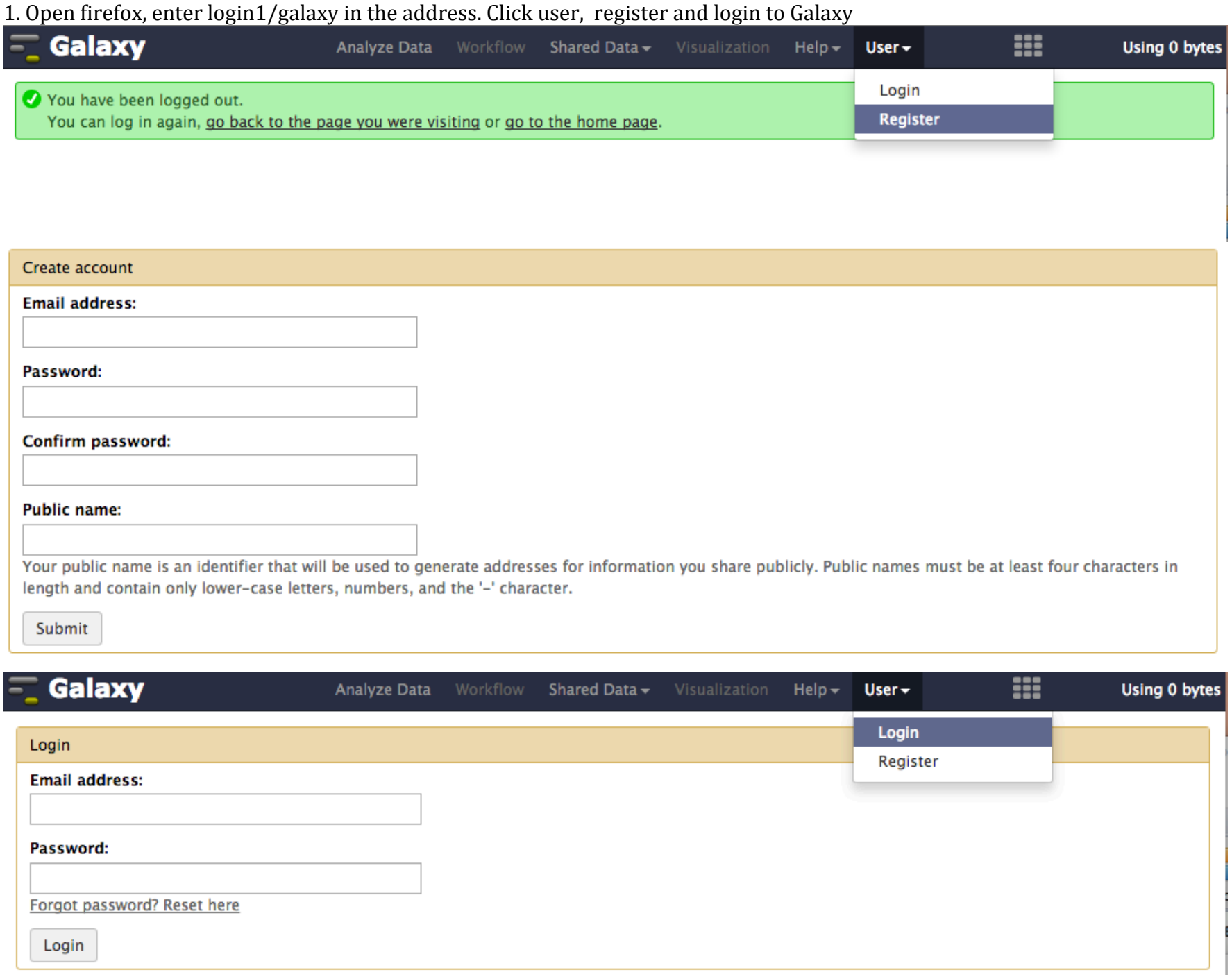
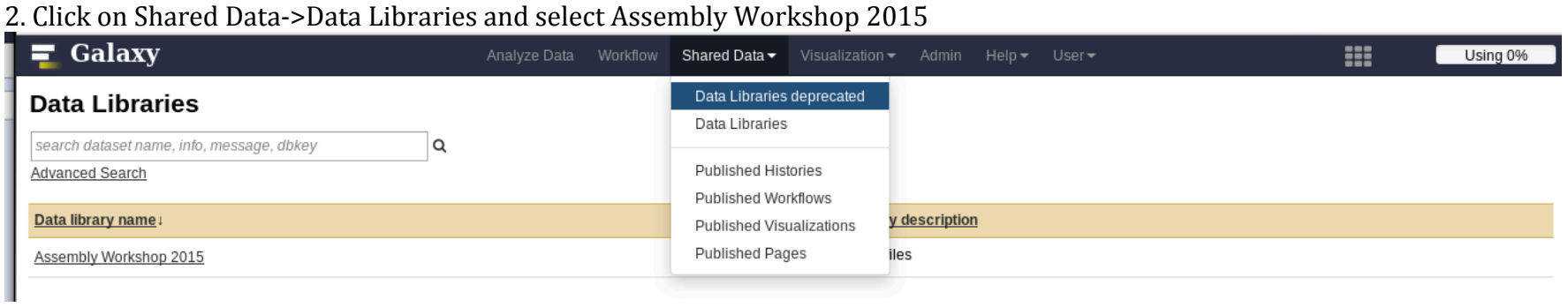
CGEBM Assembly Workshop December 2015
Galaxy Walkthrough

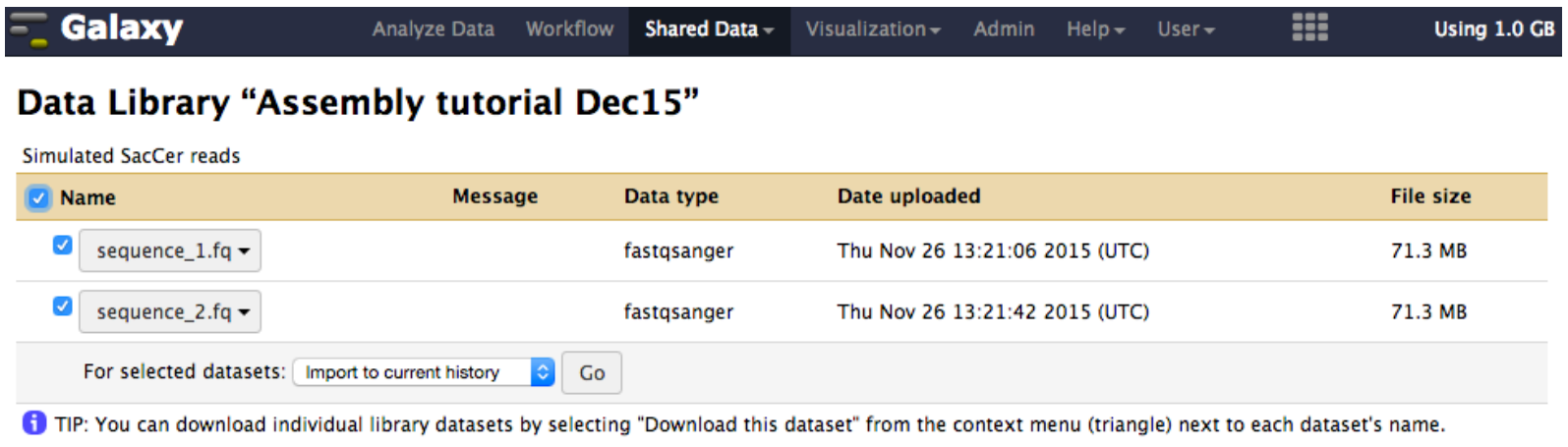
1. Open firefox, enter login1/galaxy in the address. Click user, register and login to Galaxy



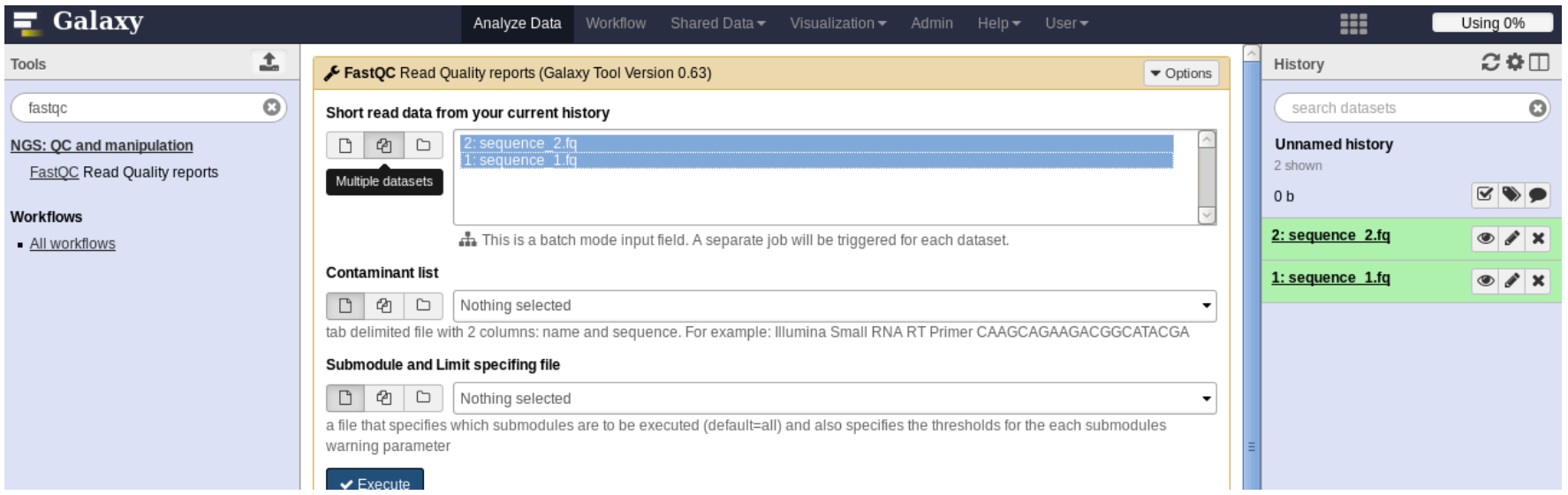
2. Click on Shared Data->Data Libraries and select Assembly Workshop 2015



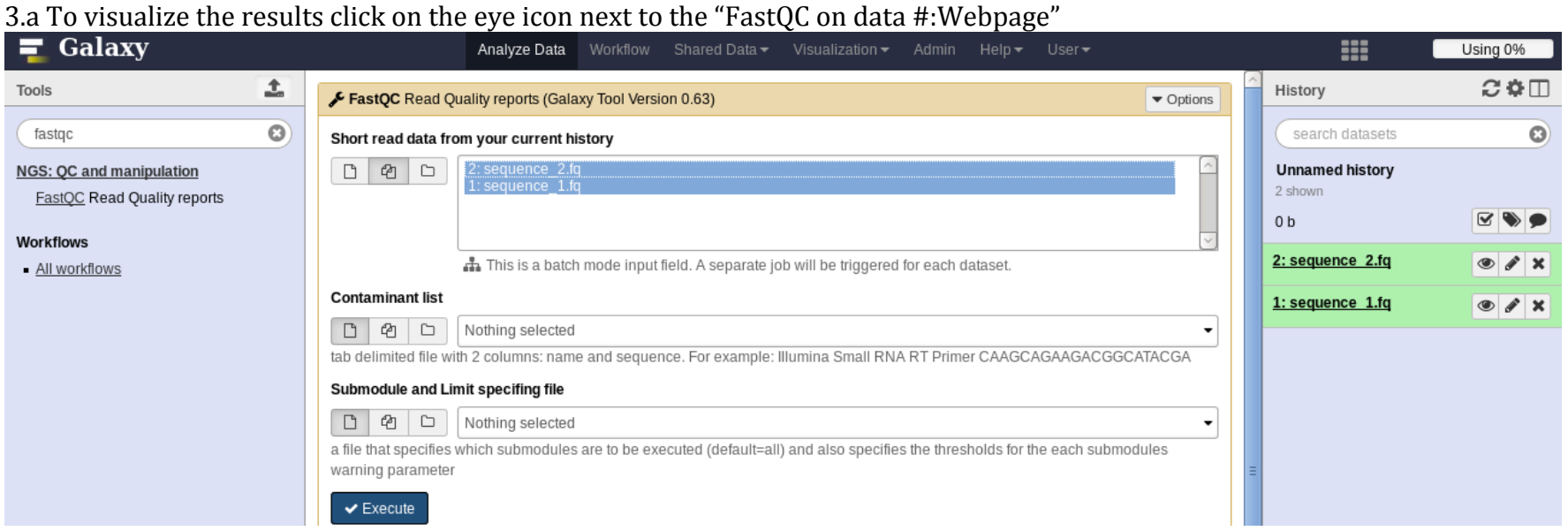
2.b Select sequence_1.fq and sequence_2.fq -> import to current history and click Go

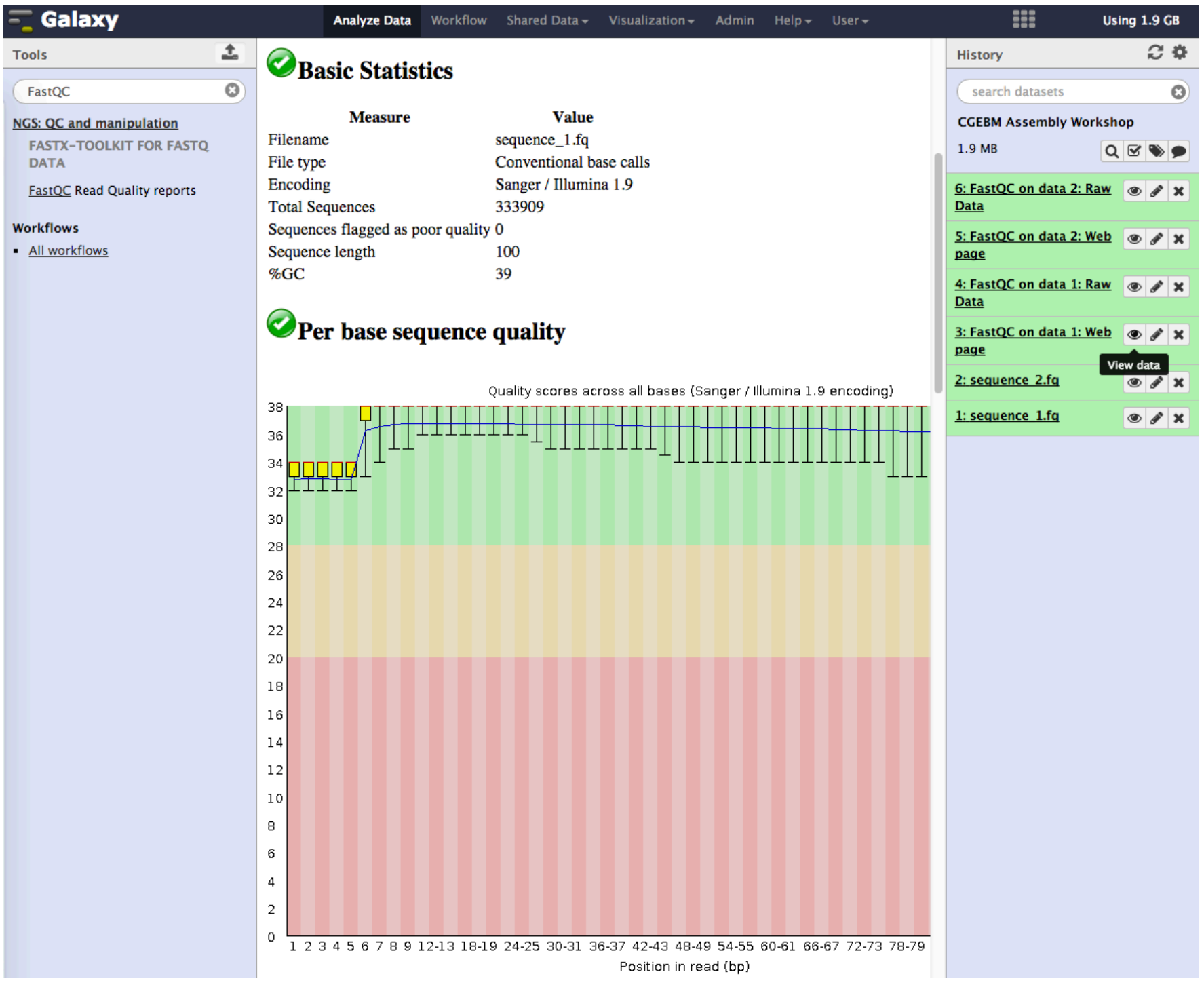


3. Select Analyze Data tab, then search and choose FastQC. Select multiple datasets, shift click to selet sequence_1.fq and sequence_2.fq and click Execute.

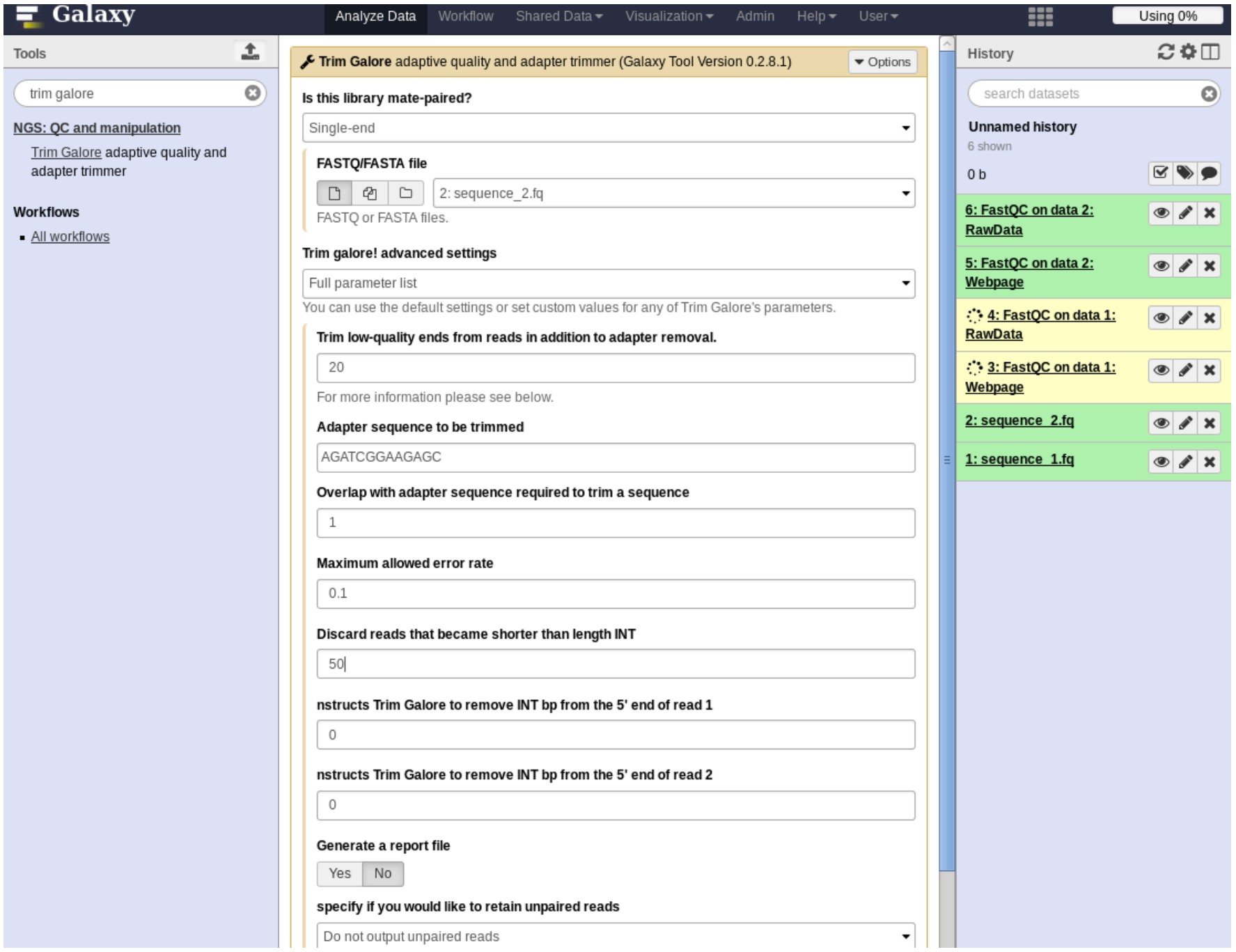


3.a To visualize the results click on the eye icon next to the “FastQC on data #:Webpage”

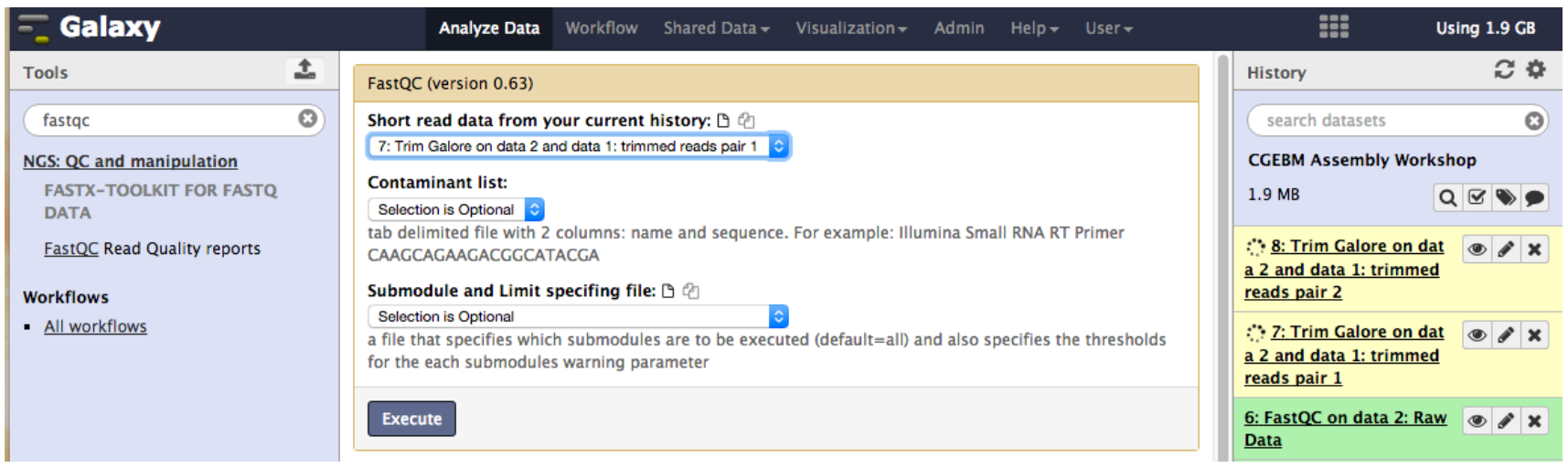




4. Search and choose Trim_Galore In the middle panel, select Paired-end and sequence_1.fq, sequence_2.fq. Set the “Discard Reads that became shorter than int” to 50 and click Execute button.

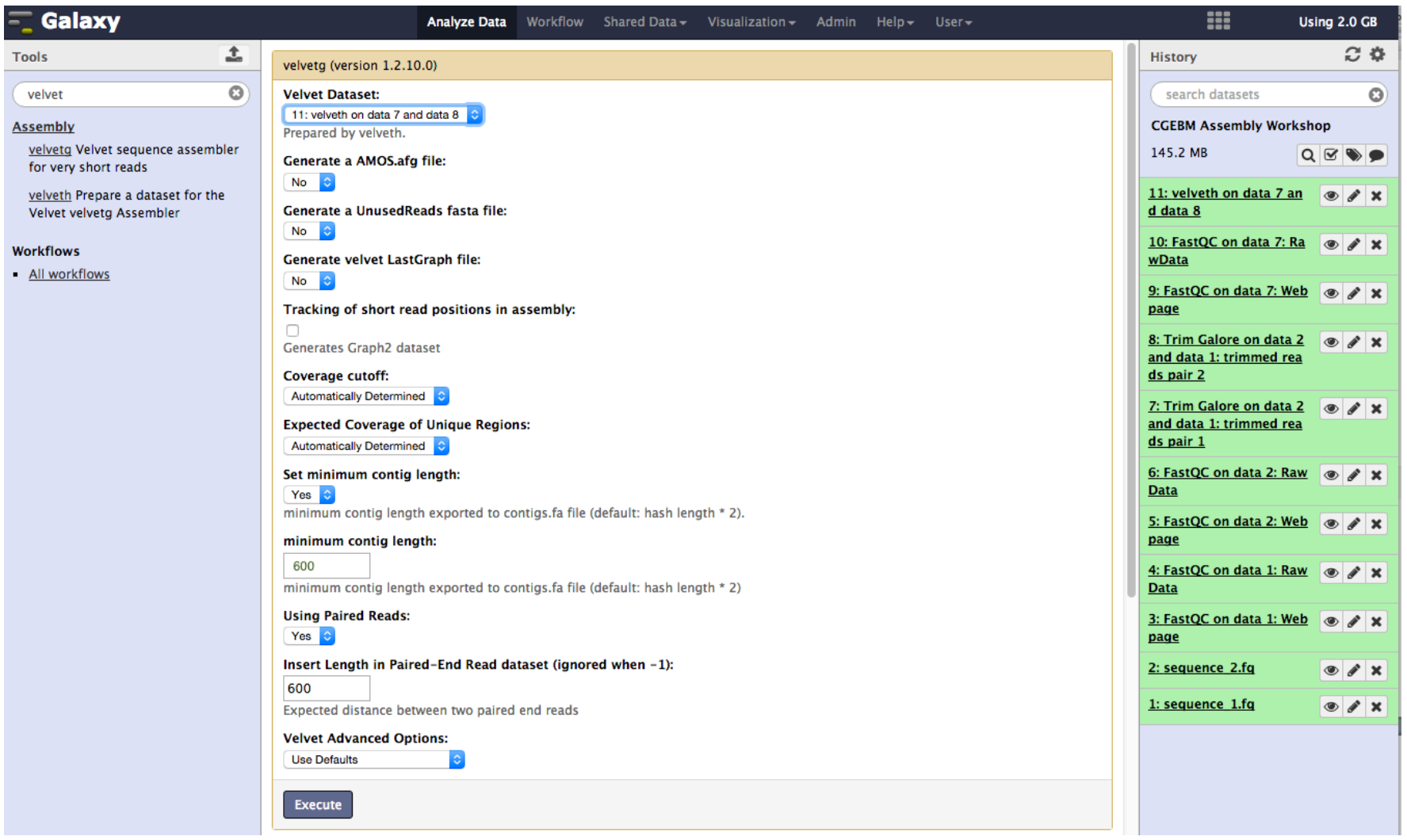


5. Repeat FastQC using Trim_Galore results

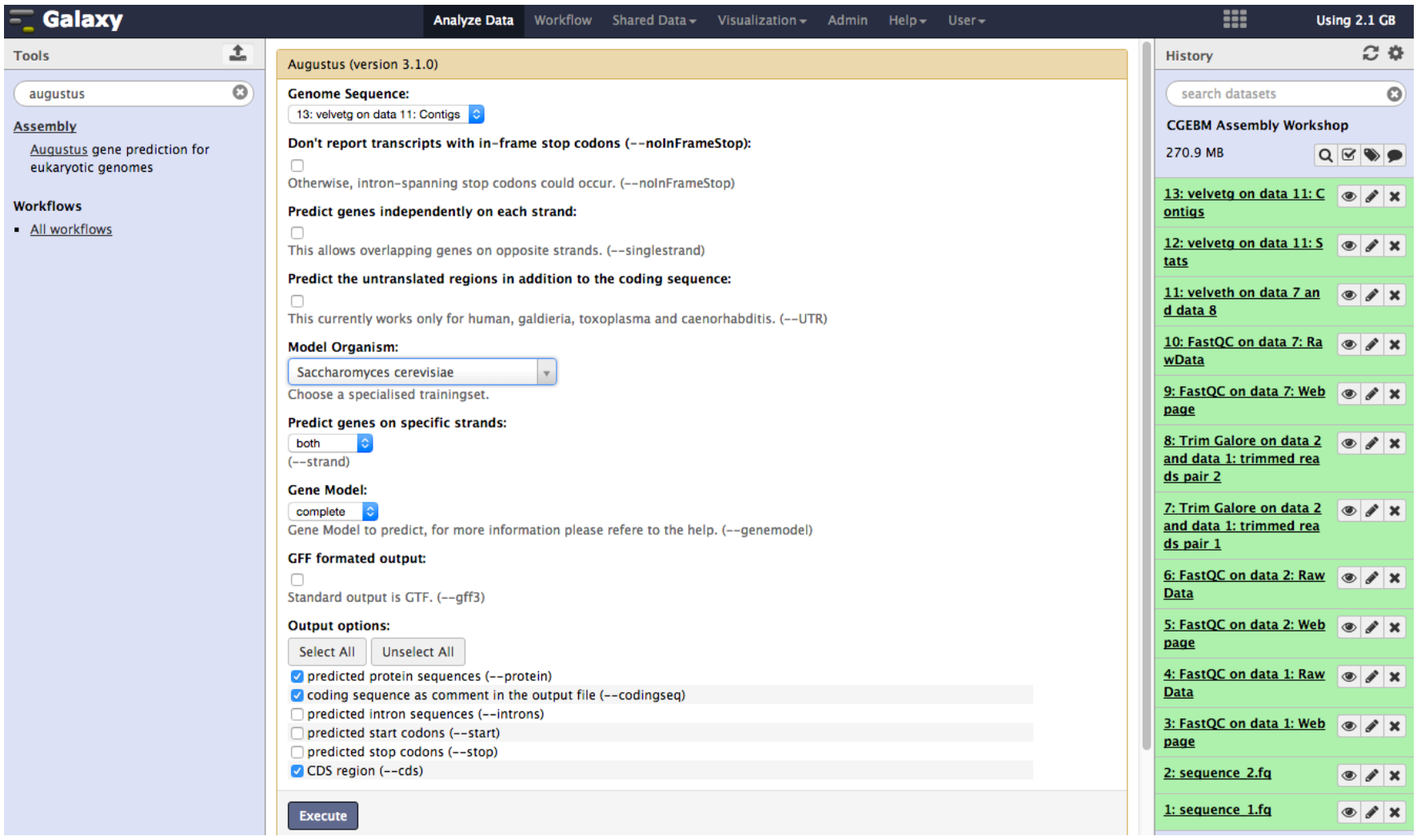


6. Search and select velveth on the Tools menu. “Add new input file”, set file to “fastq” and read type to “shortPaired reads” and Dataset “trimmed reads pair 1”. Click on “Add new input file” and file to “fastq” and read type to “shortPaired reads” and Dataset “trimmed reads pair 2”

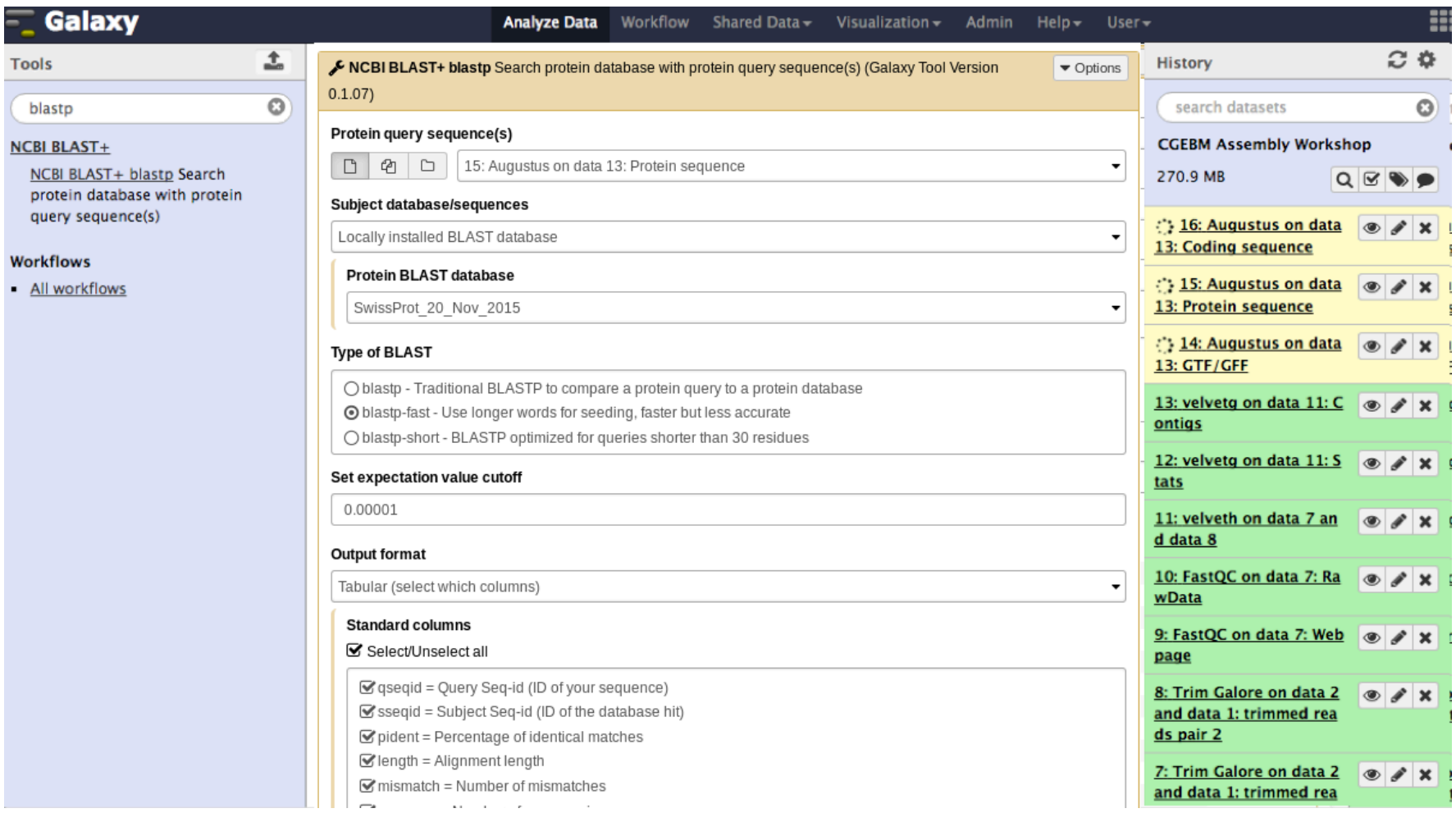
7. Search for velvetg, select the velveth results change “coverage cutoff” and “expected coverage” to “automatically determined” . Set minimum contig length to 600 and using paired reads to “Yes” with Insert Length 600. Click Execute



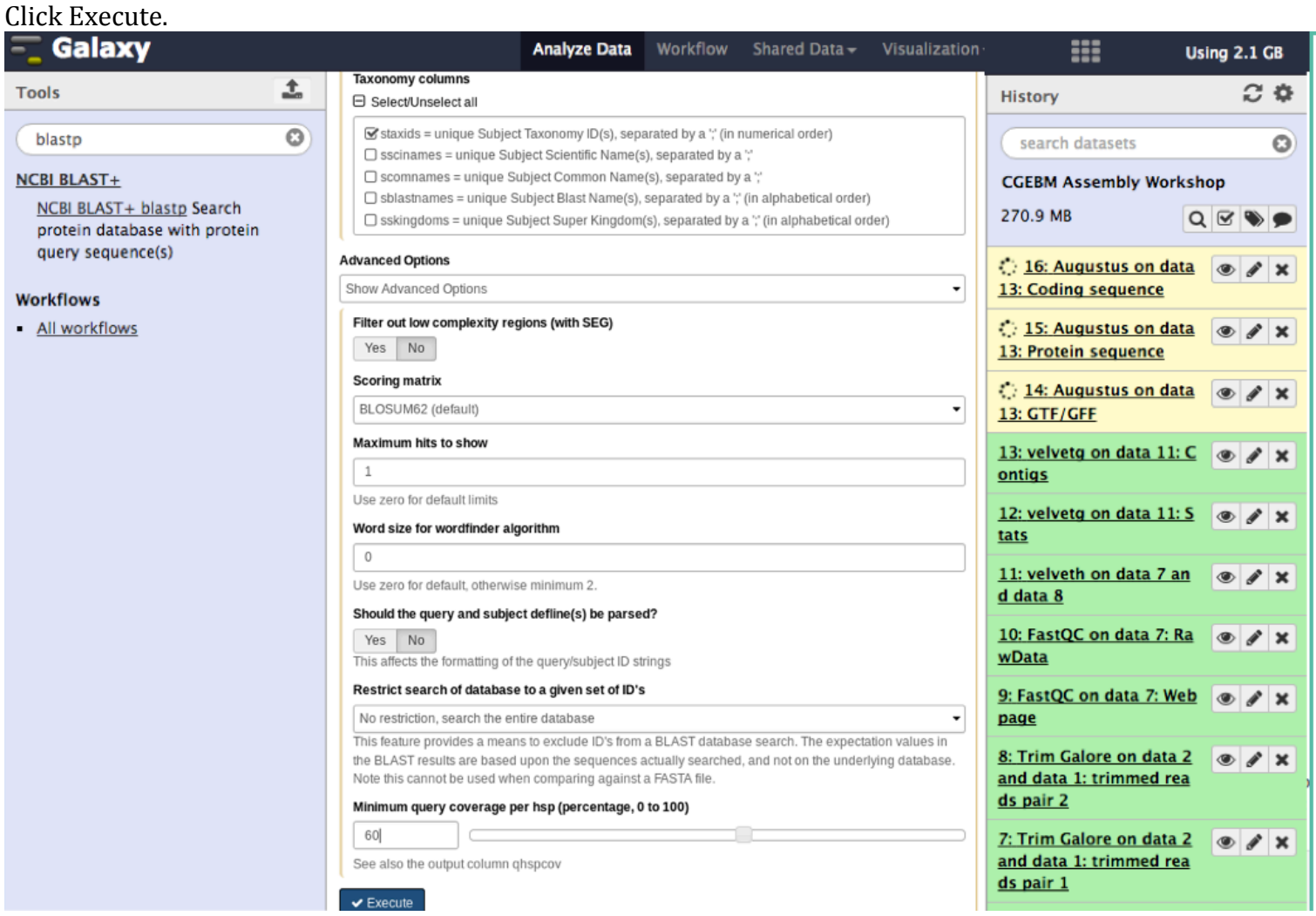
8. Search for Augustus in the Tools pane and select velvetg Contigs as input and Model Organism “Saccharomyces cerevisiae”



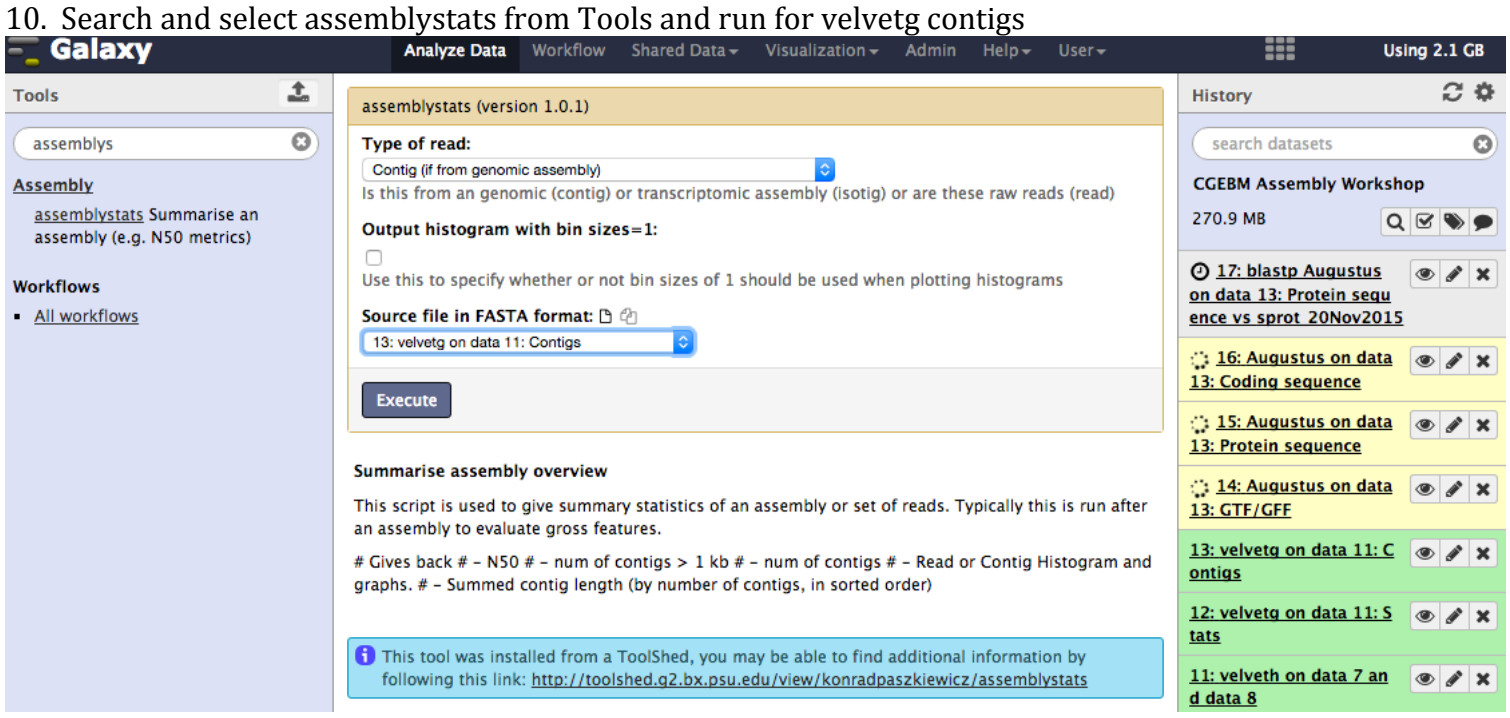
9. Search for blastp in the Tools pane, select Augustus protein sequence as query, and Swiss-Prot as database. Change the type of blast to blastp-fast, expectation cut-off to 0.00001 and click on tabular (select which columns) and scroll down to select extra columns as per the next slide



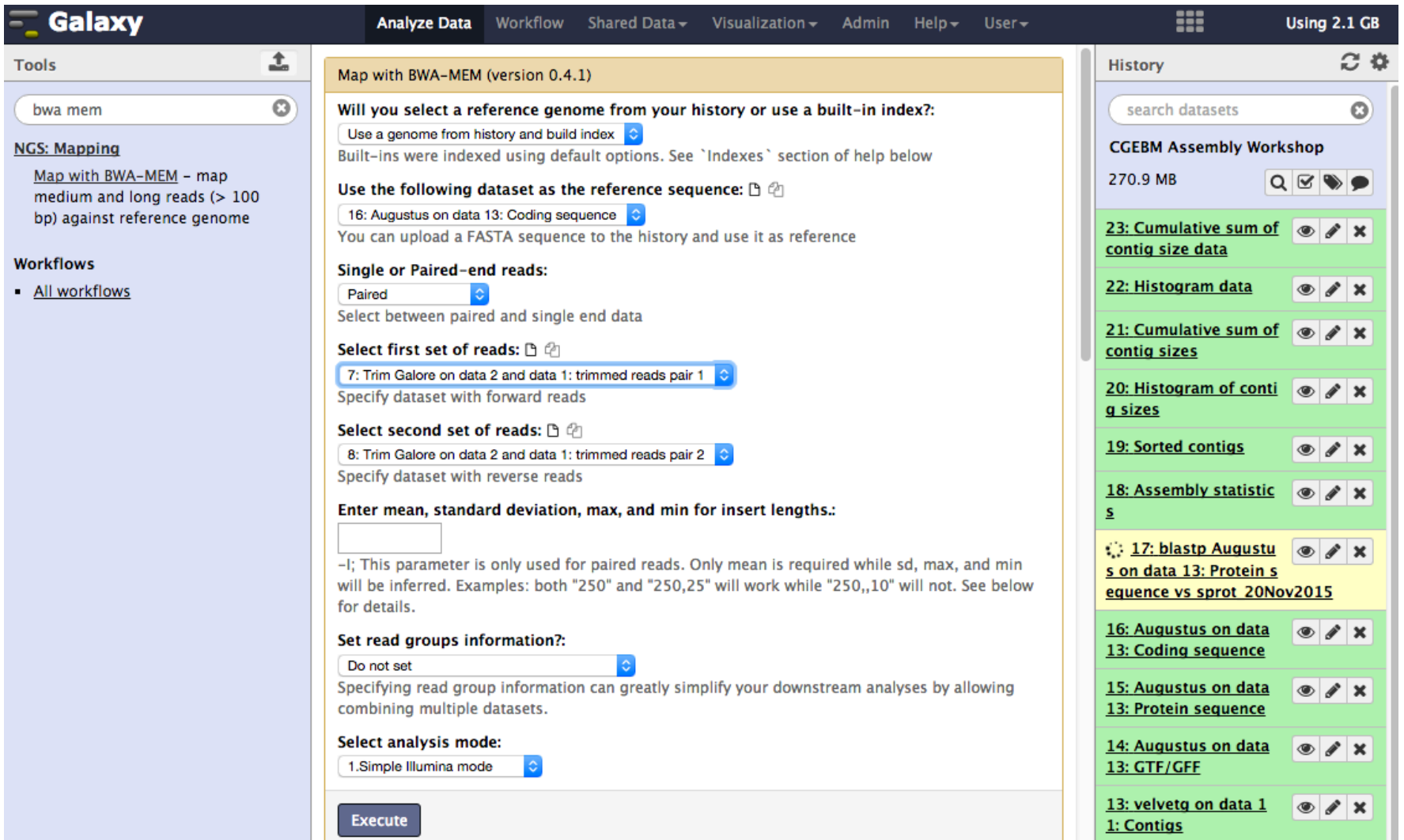
9.a Scroll and select staxids. Click on show advanced options and set maximum hits to 1 and minimum query coverage to 60. Click Execute.



10. Search and select assemblystats from Tools and run for velvetg contigs



11. Select BWA mem from Tools. “Use a genome from history” and select Augutus Coding Sequence. Select Paired reads and “trimmed reads pair 1” and “trimmed reads pair 2” . Click on Execute.



12. Select blobplot from Tools pane. Select blastp results: blastp “Augustus on Protein sequence vs sprot” , predicted Nucleotide “Augustus Coding Sequence”, Aligned reads “Map with BWA_MEM”, cut-off to 0.001 and Taxon level to Order. Click Execute.