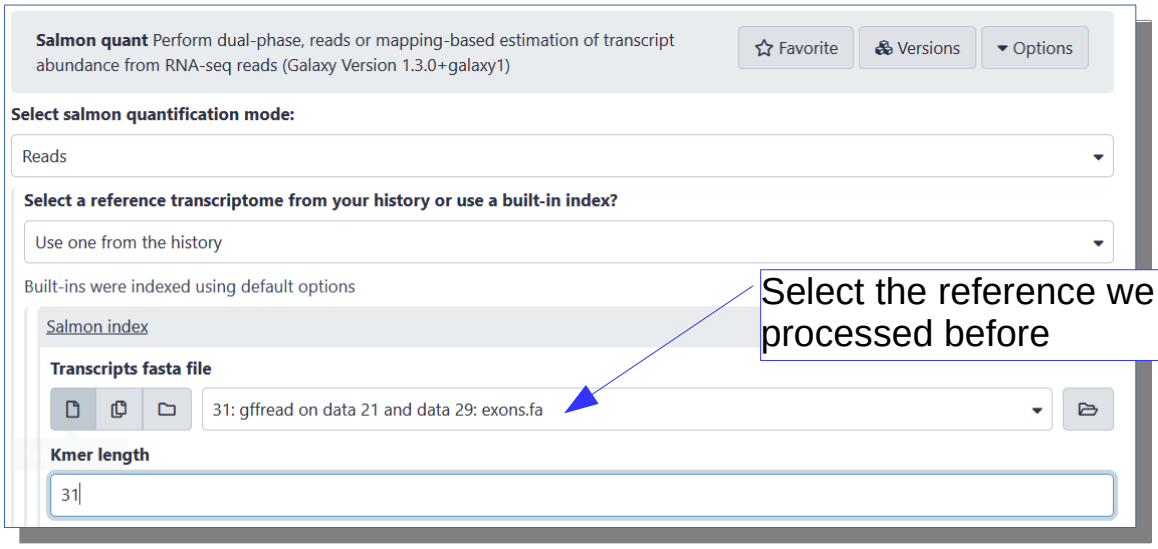
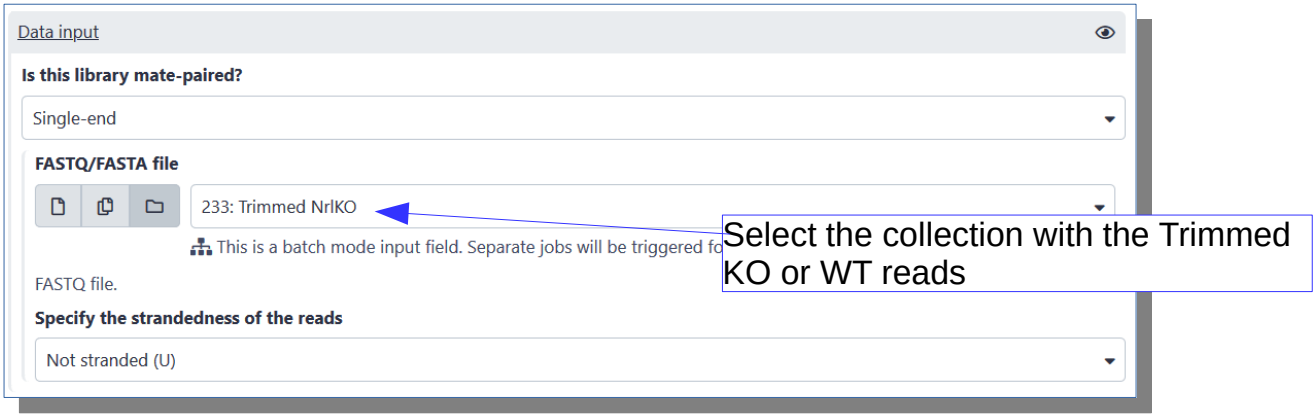
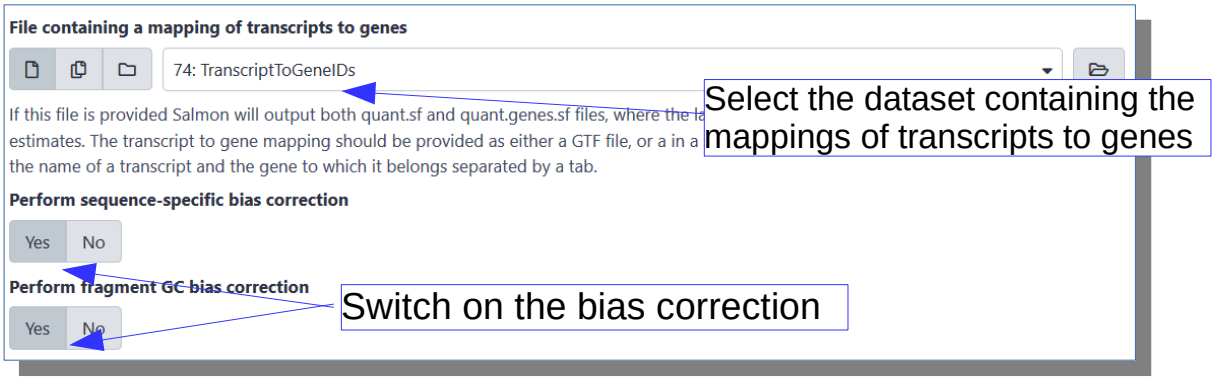
RNA-Seq Analysis: Align and Quantify

Alignment and Quantification using Salmon

- We will provide to salmon the reference, the collection of fastq files and the mapping between transcripts and genes.







- Leave the rest of the parameters as default. Usually the tools are setup with default options that suits most of the analysis problems. Do not change then unless you now what they mean.
- Repeat the process for the other samples.

RNA-Seq Analysis: Align and Quantify

- Salmon produces several outputs: Transcript and Gene quantification level.

1	2	3	4	5
Name	Length	EffectiveLength	TPM	NumReads
ENSMUST00000193812	1070	820.000	0.000000	0.000
ENSMUST00000082908	110	4.000	0.000000	0.000
ENSMUST00000162897	4153	3378.824	2.859154	340.095
ENSMUST00000159265	2989	2399.420	1.542124	130.264
ENSMUST00000070533	3634	2981.014	2.121509	222.641
ENSMUST00000195335	2819	2484.967	0.022862	2.000
ENSMUST00000192336	2233	1983.000	0.000000	0.000
ENSMUST00000194099	2309	1764.923	0.016094	1.000
ENSMUST00000161581	250	20.000	0.000000	0.000
ENSMUST00000192973	2057	1807.000	0.000000	0.000
ENSMUST00000195166	3012	2496.259	0.669989	58.878

Transcript Level Quantification

1	2	3	4	5
Name	Length	EffectiveLength	TPM	NumReads
ENSMUSG00000096768	1156.96	660.115	32.6979	759.865
ENSMUSG00000095366	595.513	350.154	9.79608	120.756
ENSMUSG00000099399	444.5	195	0	0
ENSMUSG00000096178	844	594	0	0
ENSMUSG00000102011	437	188	0	0
ENSMUSG00000100608	931	681	0	0
ENSMUSG00000101402	967.5	718	0	0
ENSMUSG00000100492	963.5	714	0	0
ENSMUSG00000100067	967	717.5	0	0
ENSMUSG00000101984	802.5	553	0	0
ENSMUSG00000099649	397	148	0	0

Gene Level Quantification

- Salmon produces two quantifications TPMs and NumReads:
  - TPMs is a normalized quantification by the total number of transcripts produced in the experiment.
  - NumReads are just the counts, how many sequences mapped to that particular transcript or gene.