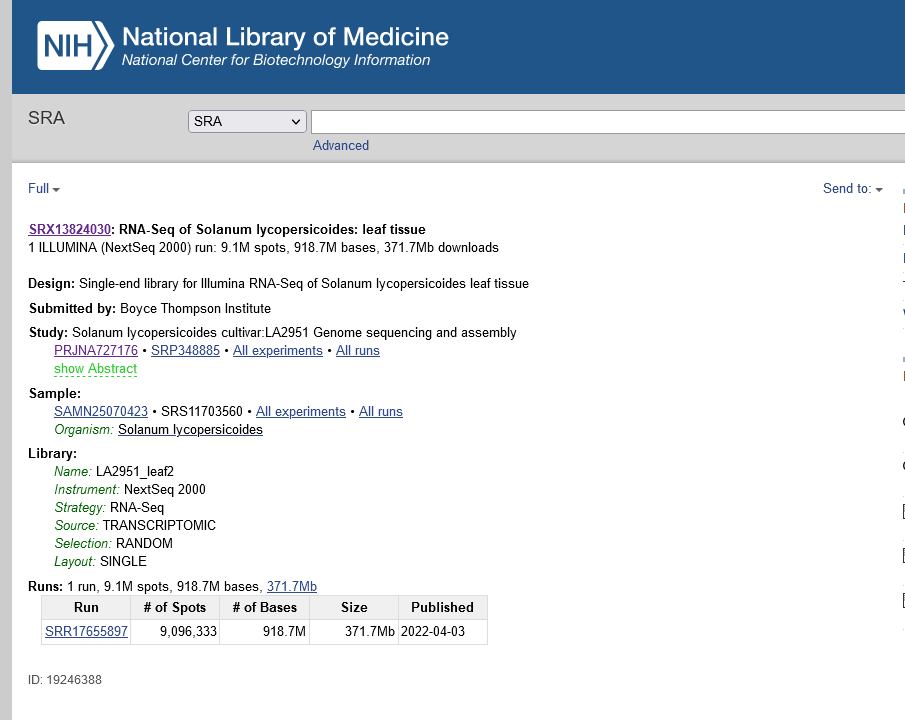
# Graded Student work

You know have a very good idea what the downloaded files were



As you see there are several different tissues. We downloaded two files per tissues so far and will process and evaluate them until the end. This means analyzing differential expression, probably a PCA, several plots generated in R and some very, very basic biological interpretation and/or hypothesis building.

After (and during the course) you have learned

* Some molecular background
* What RNASeq is its foundations and how it can be used
* Bioinformatics, processes and statistics to analyze RNAseq data
* Actual commands to work on a concrete example

Your task is to analyze in a group of maximally 3 people to analyze a tissue difference completely and document this. You can use a different tissue combination example or another project you like – ideally use a plant example though. Please only submit 1 solution per group.

If possible use the full replicate set. Depending on your machine (RAM+storage) for some of you that might be more difficult – in this case you can go down to minimally 2 replicates. Ideally though pair up with someone with a stronger computer or use appropriate short cuts.

You are free to use mapper =>counter or directly a pseudomapper. You get extra credit if you do both and perform a (technical) comparison. (up to the maximum of full credit – you can get full credit without as well (i.e. 1.0) but if you get points deducted for eg sloppy writing, missing labels etc this might help you)

A) Submit a scientific paper or report form i.e. include introduction, Materials and Methods, Results and Discussion.

Naturally your introduction can be rather technical. Using Pubmed you can find a lot of RNASeq studies as well – these however typically do not investigate 2-3 tissues differences. There you can also see how superficially things are necessarily described. Hence…

B) You also need to document your code and steps reproducibly. This is often done for real papers, either in supplements or in github repositories. Hence all steps and R code should be available and reproducible in the best case they should also be formatted to be understandable

Make sure to label all graphs and tables carefully. Captions of graphs and Tables should in theory be enough to understand the graphs. Make sure that the result text describes all figures and tables you use.

Use the spellchecking feature of your machine.

Also where possible cite appropriate literature e.g. at least for the different tools.

A discussion is generall critical and puts your results into the context of what was known and what is new now, and/or (more for Bsc/Msc etc) why things didn’t work out and what could be done.