Gobi Blockteil: ReCount

ReCount is an online resource consisting of RNA-seq gene and exon counts as well as coverage bigWig files for 2041 different studies. The recount2 resource summarizes expression data for genes, exons, exon–exon splice junctions and base-level coverage, which enables multiple downstream analyses, including testing for differential expression of potentially unannotated transcribed sequence.

Your task is to use the ReCount expression data for downstream analysis. Before you use the recount R package (`https://www.bioconductor.org/packages/devel/bioc/vignettes/recount/inst/doc/recount-quickstart.html44_DE_analysis`) you should check if the data produced by the ReCount pipeline is reproducible.

Steps:

1. Configure Podman (see podman guide)
2. Read Monorail-External Github (`https://github.com/langmead-lab/monorail-external`) & follow the necessary steps
   (i) Pull the Monorail Docker
   (ii) check that it works
3. Data Aquisition
   (i) find data that ReCount analyzed
   (ii) get the raw fastq files (SRA-toolkit might be helpful)
4. run monorail pipeline on aquired data
   (i) edit the run_recount_pump.sh script so that all HOST paths and the RECOUNT_INPUT_HOST/accession.txt file have the correct permission
   (ii) run the run_recount_pump.sh script with the correct input parameters
5. compare data from monorail pipeline to Recount data