Using Feature counts walk through

Step 1) Copy our scripts

Log in to the super computer Make a directory /scratch/User/yourusername/day6/ Copy the two scripts we have provided in /scratch/Shares/public/sread2021/scripts/day6/ Cd into the new directory

maryallen - maallen3@ip-172-31-38-192;/scratch/Users/maallen3/day6 - ssh maallen3@18.219.252.252 - 146×34
[maallen3@ip-172-31-38-192 ~]\$ mkdir -p /scratch/Users/maallen3/day6/
[maallen3@ip-172-31-38-192 ~]\$ scp /scratch/Shares/public/
genomes/ sread2021/
[maallen3@ip-172-31-38-192 ~]\$ scp /scratch/Shares/public/sread2021/
algorithms/ cookingShow/ data_files/ homework_data_files/ README.txt scripts/
[maallen3@ip-172-31-38-192 ~]\$ scp /scratch/Shares/public/sread2021/scripts/day6/* /scratch/Users/maallen3/day6/
[maallen3@ip-172-31-38-192 ~]\$ scp /scratch/Shares/public/sread2021/scripts/day6/* /scratch/Users/maallen3/day6/
[maallen3@ip-172-31-38-192 day6]\$ ls
counts featureCounts.R sr_featcounts.sbatch workshop

Step 2) Edit Our scripts

In the R script Change your working directory In the sbatch script change your email and error and output files ####Make sure the error and output directory exist before you run!!!!!!

📷 maryallen — maallen3@ip-172-31-38-192:/scratch/Users/maallen3/day6 — ssh maallen3@18.219.252.252 — 146×34 ####### Running featureCounts ######## Author: Taylor Jones
######## Here we will learn how to download a package, what metadata table is (and why it is important), ######## and run featureCounts, which counts reads over genes. ****** ####### We will want to start fresh and clear our environment. # start by clearing your console. To do this hit Ctrl+l or go to Edit-->Clear Console # clear your environment and plots by hitting the broom icon in both those cells # reset our working directory
workdir <- '/scratch/Users/username/day6/'</pre> becomes ####### Running featureCounts ####### Author: Taylor Jones ####### Here we will learn how to download a package, what metadata table is (and why it is important), ####### and run featureCounts, which counts reads over genes. ####### We will want to start fresh and clear our environment. # start by clearing your console. To do this hit Ctrl+l or go to Edit-->Clear Console # clear your environment and plots by hitting the broom icon in both those cells # reset our working directory
workdir <- '/scratch/Users/maallen3/day6/'</pre> setwd<mark>(</mark>workdir<mark>)</mark> aetwd()

like hefore, most neckanes you can install with this syntax, install neckanes("DARKACE"). Such as this neckane, install

```
    maryallen – maallen3@ip-172-31-38-192:/scratch/Users/maallen3/day6 – ssh maallen3@18.219.252.252 – 146×34
#1/bin/bash
```

```
      #SBATCH --job-name=featurecounts
      # Job name

      #SBATCH --mail-type=ALL
      # Mail events (NONE, BEGIN, END, FAIL, ALL)

      #SBATCH --mail-user=username@email.edu
      # Mail events (NONE, BEGIN, END, FAIL, ALL)

      #SBATCH --nodes=1
      # Number of cores job will run on

      #SBATCH --ntasks=4
      # Number of CPU (processers, tasks)

      #SBATCH --ntme=2:00:00
      # Time limit hrs:min:sec

      #SBATCH --partition compute
      # Job queue

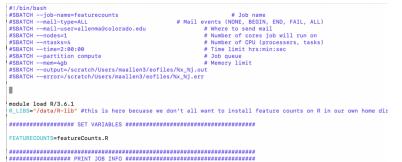
      #SBATCH --neme4gb
      # Memory limit

      #SBATCH --netry=/scratch/Users/username/eofiles/%x_%j.out
      #SBATCH --remer4gb
```

module load R/3.6.1

LIBS="/data/R-lib" #this is here becuase we don't all want to install feature counts on R in our own home directorys. BIT installed it here

becomes



Step 3) Submit the sbatch script to the queue The sbatch script runs the R script… how?

[maallen3@ip-172-31-38-192 day6]\$ sbatch sr_featcounts.sbatch Submitted batch job 1293

Step 4) If it works you will end up with a directory named counts in your working directory. In the counts directory you will end up with two files. One has the count data. One has information about the feature counts run.

```
[maallen3@ip-172-31-38-192 day6]$ ls
counts featureCounts.R sr_featcounts.sbatch
[maallen3@ip-172-31-38-192 day6]$ cd counts/
[maallen3@ip-172-31-38-192 counts]$ pwd
/scratch/Users/maallen3/day6/counts
[maallen3@ip-172-31-38-192 counts]$
```

[maallen3@ip-172-31-38-192 counts]\$ more featureCounts_gene_rnaseq.txt fc.annotation...c..GeneID... chr21Eric.repA.RNA.sorted.bam chr21Ethan.repA. RNA.sorted.bam l0C102724184 18 22 LOC102723996 1269 676 825 LOC102724023 1609 L0C102724132 10 31 59 LOC101928576 51 369 608 0 LOC102724159 866 LOC102724200 1027 LOC102724219 00102724334 100102724354 L0C102724354 L0C102724385 L0C102724370 L0C102724398 L0C102724411 1

