Getting the genomic lengths of the genes that are assigned to TADs and are found in the replicate file (i.e., GSM3449348\_RNA\_LacZ\_rep1.txt).

—----------------------------------------------------------------------------------------------------------------

(1). Extract the genes names and their genomic lengths from the Fasta file (dmel-all-gene-r5.12.fasta ) by running the command below (bioawk) in the terminal in the current folder:

bioawk -c fastx '{ print $name, length($seq) }' dmel-all-gene-r5.12.fasta >all\_gene\_genomic\_length\_5\_12.txt

The output file contains all genes and their genomic coordinates (the assigned TADs).

(2). Run “Genes\_and\_genomic\_lengths\_in\_replicate.py” script to get GeneIDs and their genomic lengths which are found in the replicate file (i.e., GSM3449348\_RNA\_LacZ\_rep1.txt) and are assigned to TADs.

(3). Run “2\_Calculate\_sum\_of\_genomic\_lengths\_in\_TADs.py” script to get the sum of the mapped genes (i.e., assigned to a TAD and are found in the replicate) in each TAD. Also, the one can run “1\_Assign\_genomic\_lengths\_to\_genes\_in\_TADs.py” script to assign the genomic lengths of the mapped genes to each TAD.

(4). If necessary, run “Convert\_txt\_to\_csv.py” script to convert the output file in TXT file from (3) to a CSV file. Output is in “Final\_Data\_for\_spreadsheet\_csv” folder.