(base) leonid@leonid-Aspire-A315-42:~$ ./FastQC/fastqc ./HW2\_sem4/ENCFF182AFY.fastq

Started analysis of ENCFF182AFY.fastq

Approx 5% complete for ENCFF182AFY.fastq

Approx 10% complete for ENCFF182AFY.fastq

Approx 15% complete for ENCFF182AFY.fastq

Approx 20% complete for ENCFF182AFY.fastq

Approx 25% complete for ENCFF182AFY.fastq

Approx 30% complete for ENCFF182AFY.fastq

Approx 35% complete for ENCFF182AFY.fastq

Approx 40% complete for ENCFF182AFY.fastq

Approx 45% complete for ENCFF182AFY.fastq

Approx 50% complete for ENCFF182AFY.fastq

Approx 55% complete for ENCFF182AFY.fastq

Approx 60% complete for ENCFF182AFY.fastq

Approx 65% complete for ENCFF182AFY.fastq

Approx 70% complete for ENCFF182AFY.fastq

Approx 75% complete for ENCFF182AFY.fastq

Approx 80% complete for ENCFF182AFY.fastq

Approx 85% complete for ENCFF182AFY.fastq

Approx 90% complete for ENCFF182AFY.fastq

Approx 95% complete for ENCFF182AFY.fastq

Approx 100% complete for ENCFF182AFY.fastq

Analysis complete for ENCFF182AFY.fastq

(base) leonid@leonid-Aspire-A315-42:~$ ./FastQC/fastqc ./HW2\_sem4/ENCFF335CQJ.fastq

Started analysis of ENCFF335CQJ.fastq

Approx 5% complete for ENCFF335CQJ.fastq

Approx 10% complete for ENCFF335CQJ.fastq

Approx 15% complete for ENCFF335CQJ.fastq

Approx 20% complete for ENCFF335CQJ.fastq

Approx 25% complete for ENCFF335CQJ.fastq

Approx 30% complete for ENCFF335CQJ.fastq

Approx 35% complete for ENCFF335CQJ.fastq

Approx 40% complete for ENCFF335CQJ.fastq

Approx 45% complete for ENCFF335CQJ.fastq

Approx 50% complete for ENCFF335CQJ.fastq

Approx 55% complete for ENCFF335CQJ.fastq

Approx 60% complete for ENCFF335CQJ.fastq

Approx 65% complete for ENCFF335CQJ.fastq

Approx 70% complete for ENCFF335CQJ.fastq

Approx 75% complete for ENCFF335CQJ.fastq

Approx 80% complete for ENCFF335CQJ.fastq

Approx 85% complete for ENCFF335CQJ.fastq

Approx 90% complete for ENCFF335CQJ.fastq

Approx 95% complete for ENCFF335CQJ.fastq

Analysis complete for ENCFF335CQJ.fastq

(base) leonid@leonid-Aspire-A315-42:~$ ./FastQC/fastqc ./HW2\_sem4/ENCFF569WIS.fastq

Started analysis of ENCFF569WIS.fastq

Approx 5% complete for ENCFF569WIS.fastq

Approx 10% complete for ENCFF569WIS.fastq

Approx 15% complete for ENCFF569WIS.fastq

Approx 20% complete for ENCFF569WIS.fastq

Approx 25% complete for ENCFF569WIS.fastq

Approx 30% complete for ENCFF569WIS.fastq

Approx 35% complete for ENCFF569WIS.fastq

Approx 40% complete for ENCFF569WIS.fastq

Approx 45% complete for ENCFF569WIS.fastq

Approx 50% complete for ENCFF569WIS.fastq

Approx 55% complete for ENCFF569WIS.fastq

Approx 60% complete for ENCFF569WIS.fastq

Approx 65% complete for ENCFF569WIS.fastq

Approx 70% complete for ENCFF569WIS.fastq

Approx 75% complete for ENCFF569WIS.fastq

Approx 80% complete for ENCFF569WIS.fastq

Approx 85% complete for ENCFF569WIS.fastq

Approx 90% complete for ENCFF569WIS.fastq

Approx 95% complete for ENCFF569WIS.fastq

Analysis complete for ENCFF569WIS.fastq

(base) leonid@leonid-Aspire-A315-42:~$ cd HW2\_sem4

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ wget https://hgdownload.soe.ucsc.edu/goldenPath/hg38/chromosomes/chr14.fa.gz

--2022-03-13 16:01:52-- https://hgdownload.soe.ucsc.edu/goldenPath/hg38/chromosomes/chr14.fa.gz

Resolving hgdownload.soe.ucsc.edu (hgdownload.soe.ucsc.edu)... 128.114.119.163

Connecting to hgdownload.soe.ucsc.edu (hgdownload.soe.ucsc.edu)|128.114.119.163|:443... connected.

HTTP request sent, awaiting response... 200 OK

Length: 29295890 (28M) [application/x-gzip]

Saving to: ‘chr14.fa.gz’

chr14.fa.gz 100%[===================================================================================>] 27,94M 53,8KB/s in 9m 9s

2022-03-13 16:11:03 (52,1 KB/s) - ‘chr14.fa.gz’ saved [29295890/29295890]

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ gunzip chr14.fa.gz

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ bowtie2-build chr14.fa chromosome\_index

Settings:

Output files: "chromosome\_index.\*.bt2"

Line rate: 6 (line is 64 bytes)

Lines per side: 1 (side is 64 bytes)

Offset rate: 4 (one in 16)

FTable chars: 10

Strings: unpacked

Max bucket size: default

Max bucket size, sqrt multiplier: default

Max bucket size, len divisor: 4

Difference-cover sample period: 1024

Endianness: little

Actual local endianness: little

Sanity checking: disabled

Assertions: disabled

Random seed: 0

Sizeofs: void\*:8, int:4, long:8, size\_t:8

Input files DNA, FASTA:

chr14.fa

Building a SMALL index

Reading reference sizes

Time reading reference sizes: 00:00:00

Calculating joined length

Writing header

Reserving space for joined string

Joining reference sequences

Time to join reference sequences: 00:00:01

bmax according to bmaxDivN setting: 22642037

Using parameters --bmax 16981528 --dcv 1024

Doing ahead-of-time memory usage test

Passed! Constructing with these parameters: --bmax 16981528 --dcv 1024

Constructing suffix-array element generator

Building DifferenceCoverSample

Building sPrime

Building sPrimeOrder

V-Sorting samples

V-Sorting samples time: 00:00:01

Allocating rank array

Ranking v-sort output

Ranking v-sort output time: 00:00:01

Invoking Larsson-Sadakane on ranks

Invoking Larsson-Sadakane on ranks time: 00:00:00

Sanity-checking and returning

Building samples

Reserving space for 12 sample suffixes

Generating random suffixes

QSorting 12 sample offsets, eliminating duplicates

QSorting sample offsets, eliminating duplicates time: 00:00:00

Multikey QSorting 12 samples

(Using difference cover)

Multikey QSorting samples time: 00:00:00

Calculating bucket sizes

Splitting and merging

Splitting and merging time: 00:00:00

Avg bucket size: 9.05681e+07 (target: 16981527)

Converting suffix-array elements to index image

Allocating ftab, absorbFtab

Entering Ebwt loop

Getting block 1 of 1

No samples; assembling all-inclusive block

Sorting block of length 90568149 for bucket 1

(Using difference cover)

Sorting block time: 00:00:19

Returning block of 90568150 for bucket 1

Exited Ebwt loop

fchr[A]: 0

fchr[C]: 26673415

fchr[G]: 45097173

fchr[T]: 63656206

fchr[$]: 90568149

Exiting Ebwt::buildToDisk()

Returning from initFromVector

Wrote 34384187 bytes to primary EBWT file: chromosome\_index.1.bt2

Wrote 22642044 bytes to secondary EBWT file: chromosome\_index.2.bt2

Re-opening \_in1 and \_in2 as input streams

Returning from Ebwt constructor

Headers:

len: 90568149

bwtLen: 90568150

sz: 22642038

bwtSz: 22642038

lineRate: 6

offRate: 4

offMask: 0xfffffff0

ftabChars: 10

eftabLen: 20

eftabSz: 80

ftabLen: 1048577

ftabSz: 4194308

offsLen: 5660510

offsSz: 22642040

lineSz: 64

sideSz: 64

sideBwtSz: 48

sideBwtLen: 192

numSides: 471710

numLines: 471710

ebwtTotLen: 30189440

ebwtTotSz: 30189440

color: 0

reverse: 0

Total time for call to driver() for forward index: 00:00:33

Reading reference sizes

Time reading reference sizes: 00:00:01

Calculating joined length

Writing header

Reserving space for joined string

Joining reference sequences

Time to join reference sequences: 00:00:00

Time to reverse reference sequence: 00:00:01

bmax according to bmaxDivN setting: 22642037

Using parameters --bmax 16981528 --dcv 1024

Doing ahead-of-time memory usage test

Passed! Constructing with these parameters: --bmax 16981528 --dcv 1024

Constructing suffix-array element generator

Building DifferenceCoverSample

Building sPrime

Building sPrimeOrder

V-Sorting samples

V-Sorting samples time: 00:00:01

Allocating rank array

Ranking v-sort output

Ranking v-sort output time: 00:00:00

Invoking Larsson-Sadakane on ranks

Invoking Larsson-Sadakane on ranks time: 00:00:01

Sanity-checking and returning

Building samples

Reserving space for 12 sample suffixes

Generating random suffixes

QSorting 12 sample offsets, eliminating duplicates

QSorting sample offsets, eliminating duplicates time: 00:00:00

Multikey QSorting 12 samples

(Using difference cover)

Multikey QSorting samples time: 00:00:00

Calculating bucket sizes

Splitting and merging

Splitting and merging time: 00:00:00

Avg bucket size: 9.05681e+07 (target: 16981527)

Converting suffix-array elements to index image

Allocating ftab, absorbFtab

Entering Ebwt loop

Getting block 1 of 1

No samples; assembling all-inclusive block

Sorting block of length 90568149 for bucket 1

(Using difference cover)

Sorting block time: 00:00:19

Returning block of 90568150 for bucket 1

Exited Ebwt loop

fchr[A]: 0

fchr[C]: 26673415

fchr[G]: 45097173

fchr[T]: 63656206

fchr[$]: 90568149

Exiting Ebwt::buildToDisk()

Returning from initFromVector

Wrote 34384187 bytes to primary EBWT file: chromosome\_index.rev.1.bt2

Wrote 22642044 bytes to secondary EBWT file: chromosome\_index.rev.2.bt2

Re-opening \_in1 and \_in2 as input streams

Returning from Ebwt constructor

Headers:

len: 90568149

bwtLen: 90568150

sz: 22642038

bwtSz: 22642038

lineRate: 6

offRate: 4

offMask: 0xfffffff0

ftabChars: 10

eftabLen: 20

eftabSz: 80

ftabLen: 1048577

ftabSz: 4194308

offsLen: 5660510

offsSz: 22642040

lineSz: 64

sideSz: 64

sideBwtSz: 48

sideBwtLen: 192

numSides: 471710

numLines: 471710

ebwtTotLen: 30189440

ebwtTotSz: 30189440

color: 0

reverse: 1

Total time for backward call to driver() for mirror index: 00:00:33

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ mkdir bowtie2\_res

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ bowtie2 -p 2 -x chromosome\_index -U ENCFF182AFY.fastq -S bowtie2\_res/res\_AFY.sam

36622019 reads; of these:

36622019 (100.00%) were unpaired; of these:

32759520 (89.45%) aligned 0 times

1453730 (3.97%) aligned exactly 1 time

2408769 (6.58%) aligned >1 times

10.55% overall alignment rate

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ bowtie2 -p 2 -x chromosome\_index -U ENCFF335CQJ.fastq -S bowtie2\_res/res\_CQJ.sam

37314930 reads; of these:

37314930 (100.00%) were unpaired; of these:

33216927 (89.02%) aligned 0 times

1496129 (4.01%) aligned exactly 1 time

2601874 (6.97%) aligned >1 times

10.98% overall alignment rate

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ bowtie2 -p 2 -x chromosome\_index -U ENCFF569WIS.fastq -S bowtie2\_res/res\_WIS.sam

47833064 reads; of these:

47833064 (100.00%) were unpaired; of these:

41590140 (86.95%) aligned 0 times

2039214 (4.26%) aligned exactly 1 time

4203710 (8.79%) aligned >1 times

13.05% overall alignment rate

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ mkdir macs2

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ macs2 --help

usage: macs2 [-h] [--version] {callpeak,bdgpeakcall,bdgbroadcall,bdgcmp,bdgopt,cmbreps,bdgdiff,filterdup,predictd,pileup,randsample,refinepeak} ...

macs2 -- Model-based Analysis for ChIP-Sequencing

positional arguments:

{callpeak,bdgpeakcall,bdgbroadcall,bdgcmp,bdgopt,cmbreps,bdgdiff,filterdup,predictd,pileup,randsample,refinepeak}

callpeak Main MACS2 Function: Call peaks from alignment results.

bdgpeakcall Call peaks from bedGraph output. Note: All regions on the same chromosome in the bedGraph file should be continuous so only bedGraph files

from MACS2 are accpetable.

bdgbroadcall Call broad peaks from bedGraph output. Note: All regions on the same chromosome in the bedGraph file should be continuous so only bedGraph

files from MACS2 are accpetable.

bdgcmp Deduct noise by comparing two signal tracks in bedGraph. Note: All regions on the same chromosome in the bedGraph file should be continuous

so only bedGraph files from MACS2 are accpetable.

bdgopt Operations on score column of bedGraph file. Note: All regions on the same chromosome in the bedGraph file should be continuous so only

bedGraph files from MACS2 are accpetable.

cmbreps Combine BEDGraphs of scores from replicates. Note: All regions on the same chromosome in the bedGraph file should be continuous so only

bedGraph files from MACS2 are accpetable.

bdgdiff Differential peak detection based on paired four bedgraph files. Note: All regions on the same chromosome in the bedGraph file should be

continuous so only bedGraph files from MACS2 are accpetable.

filterdup Remove duplicate reads at the same position, then save the rest alignments to BED or BEDPE file. If you use '--keep-dup all option', this

script can be utilized to convert any acceptable format into BED or BEDPE format.

predictd Predict d or fragment size from alignment results. \*Will NOT filter duplicates\*

pileup Pileup aligned reads with a given extension size (fragment size or d in MACS language). Note there will be no step for duplicate reads

filtering or sequencing depth scaling, so you may need to do certain pre/post-processing.

randsample Randomly sample number/percentage of total reads.

refinepeak (Experimental) Take raw reads alignment, refine peak summits and give scores measuring balance of waston/crick tags. Inspired by SPP.

optional arguments:

-h, --help show this help message and exit

--version show program's version number and exit

For command line options of each command, type: macs2 COMMAND -h

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ macs2 callpeak --broad -t bowtie2\_res/res\_AFY.sam -c bowtie2\_res/res\_WIS.sam -f SAM --outdir macs2

Traceback (most recent call last):

File "/home/leonid/anaconda3/bin/macs2", line 653, in <module>

main()

File "/home/leonid/anaconda3/bin/macs2", line 49, in main

from MACS2.callpeak\_cmd import run

File "/home/leonid/anaconda3/lib/python3.8/site-packages/MACS2/callpeak\_cmd.py", line 23, in <module>

from MACS2.OptValidator import opt\_validate

File "/home/leonid/anaconda3/lib/python3.8/site-packages/MACS2/OptValidator.py", line 20, in <module>

from MACS2.IO.Parser import BEDParser, ELANDResultParser, ELANDMultiParser, \

File "\_\_init\_\_.pxd", line 242, in init MACS2.IO.Parser

ValueError: numpy.ndarray size changed, may indicate binary incompatibility. Expected 96 from C header, got 88 from PyObject

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ pip install --upgrade numpy

Requirement already satisfied: numpy in /home/leonid/anaconda3/lib/python3.8/site-packages (1.21.4)

Collecting numpy

Using cached numpy-1.22.3-cp38-cp38-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (16.8 MB)

Installing collected packages: numpy

Attempting uninstall: numpy

Found existing installation: numpy 1.21.4

Uninstalling numpy-1.21.4:

Successfully uninstalled numpy-1.21.4

Successfully installed numpy-1.22.3

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ macs2 callpeak --broad -t bowtie2\_res/res\_AFY.sam -c bowtie2\_res/res\_WIS.sam -f SAM --outdir macs2

INFO @ Sun, 13 Mar 2022 19:14:45:

# Command line: callpeak --broad -t bowtie2\_res/res\_AFY.sam -c bowtie2\_res/res\_WIS.sam -f SAM --outdir macs2

# ARGUMENTS LIST:

# name = NA

# format = SAM

# ChIP-seq file = ['bowtie2\_res/res\_AFY.sam']

# control file = ['bowtie2\_res/res\_WIS.sam']

# effective genome size = 2.70e+09

# band width = 300

# model fold = [5, 50]

# qvalue cutoff for narrow/strong regions = 5.00e-02

# qvalue cutoff for broad/weak regions = 1.00e-01

# The maximum gap between significant sites is assigned as the read length/tag size.

# The minimum length of peaks is assigned as the predicted fragment length "d".

# Larger dataset will be scaled towards smaller dataset.

# Range for calculating regional lambda is: 1000 bps and 10000 bps

# Broad region calling is on

# Paired-End mode is off

INFO @ Sun, 13 Mar 2022 19:14:45: #1 read tag files...

INFO @ Sun, 13 Mar 2022 19:14:45: #1 read treatment tags...

INFO @ Sun, 13 Mar 2022 19:14:47: 2000000

INFO @ Sun, 13 Mar 2022 19:15:02: 18000000

INFO @ Sun, 13 Mar 2022 19:15:03: 19000000

INFO @ Sun, 13 Mar 2022 19:15:13: 30000000

INFO @ Sun, 13 Mar 2022 19:15:18: 36000000

INFO @ Sun, 13 Mar 2022 19:15:19: #1.2 read input tags...

INFO @ Sun, 13 Mar 2022 19:15:22: 3000000

INFO @ Sun, 13 Mar 2022 19:15:26: 7000000

INFO @ Sun, 13 Mar 2022 19:15:42: 23000000

INFO @ Sun, 13 Mar 2022 19:15:45: 26000000

INFO @ Sun, 13 Mar 2022 19:16:04: 45000000

INFO @ Sun, 13 Mar 2022 19:16:07: #1 tag size is determined as 76 bps

INFO @ Sun, 13 Mar 2022 19:16:07: #1 tag size = 76.0

INFO @ Sun, 13 Mar 2022 19:16:07: #1 total tags in treatment: 3862499

INFO @ Sun, 13 Mar 2022 19:16:07: #1 user defined the maximum tags...

INFO @ Sun, 13 Mar 2022 19:16:07: #1 filter out redundant tags at the same location and the same strand by allowing at most 1 tag(s)

INFO @ Sun, 13 Mar 2022 19:16:07: #1 tags after filtering in treatment: 3089324

INFO @ Sun, 13 Mar 2022 19:16:07: #1 Redundant rate of treatment: 0.20

INFO @ Sun, 13 Mar 2022 19:16:07: #1 total tags in control: 6242924

INFO @ Sun, 13 Mar 2022 19:16:07: #1 user defined the maximum tags...

INFO @ Sun, 13 Mar 2022 19:16:07: #1 filter out redundant tags at the same location and the same strand by allowing at most 1 tag(s)

INFO @ Sun, 13 Mar 2022 19:16:07: #1 tags after filtering in control: 4676276

INFO @ Sun, 13 Mar 2022 19:16:07: #1 Redundant rate of control: 0.25

INFO @ Sun, 13 Mar 2022 19:16:07: #1 finished!

INFO @ Sun, 13 Mar 2022 19:16:07: #2 Build Peak Model...

INFO @ Sun, 13 Mar 2022 19:16:07: #2 looking for paired plus/minus strand peaks...

INFO @ Sun, 13 Mar 2022 19:16:09: #2 number of paired peaks: 37584

INFO @ Sun, 13 Mar 2022 19:16:09: start model\_add\_line...

INFO @ Sun, 13 Mar 2022 19:16:09: start X-correlation...

INFO @ Sun, 13 Mar 2022 19:16:09: end of X-cor

INFO @ Sun, 13 Mar 2022 19:16:09: #2 finished!

INFO @ Sun, 13 Mar 2022 19:16:09: #2 predicted fragment length is 75 bps

INFO @ Sun, 13 Mar 2022 19:16:09: #2 alternative fragment length(s) may be 75,460 bps

INFO @ Sun, 13 Mar 2022 19:16:09: #2.2 Generate R script for model : macs2/NA\_model.r

WARNING @ Sun, 13 Mar 2022 19:16:09: #2 Since the d (75) calculated from paired-peaks are smaller than 2\*tag length, it may be influenced by unknown sequencing problem!

WARNING @ Sun, 13 Mar 2022 19:16:09: #2 You may need to consider one of the other alternative d(s): 75,460

WARNING @ Sun, 13 Mar 2022 19:16:09: #2 You can restart the process with --nomodel --extsize XXX with your choice or an arbitrary number. Nontheless, MACS will continute computing.

INFO @ Sun, 13 Mar 2022 19:16:09: #3 Call peaks...

INFO @ Sun, 13 Mar 2022 19:16:09: #3 Call broad peaks with given level1 -log10qvalue cutoff and level2: 1.301030, 1.000000...

INFO @ Sun, 13 Mar 2022 19:16:09: #3 Pre-compute pvalue-qvalue table...

INFO @ Sun, 13 Mar 2022 19:16:30: #3 Call peaks for each chromosome...

INFO @ Sun, 13 Mar 2022 19:16:42: #4 Write output xls file... macs2/NA\_peaks.xls

INFO @ Sun, 13 Mar 2022 19:16:42: #4 Write broad peak in broadPeak format file... macs2/NA\_peaks.broadPeak

INFO @ Sun, 13 Mar 2022 19:16:42: #4 Write broad peak in bed12/gappedPeak format file... macs2/NA\_peaks.gappedPeak

INFO @ Sun, 13 Mar 2022 19:16:42: Done!

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ macs2 callpeak --broad -t bowtie2\_res/res\_CQJ.sam -c bowtie2\_res/res\_WIS.sam -f SAM --outdir macs2

INFO @ Sun, 13 Mar 2022 19:18:00:

# Command line: callpeak --broad -t bowtie2\_res/res\_CQJ.sam -c bowtie2\_res/res\_WIS.sam -f SAM --outdir macs2

# ARGUMENTS LIST:

# name = NA

# format = SAM

# ChIP-seq file = ['bowtie2\_res/res\_CQJ.sam']

# control file = ['bowtie2\_res/res\_WIS.sam']

# effective genome size = 2.70e+09

# band width = 300

# model fold = [5, 50]

# qvalue cutoff for narrow/strong regions = 5.00e-02

# qvalue cutoff for broad/weak regions = 1.00e-01

# The maximum gap between significant sites is assigned as the read length/tag size.

# The minimum length of peaks is assigned as the predicted fragment length "d".

# Larger dataset will be scaled towards smaller dataset.

# Range for calculating regional lambda is: 1000 bps and 10000 bps

# Broad region calling is on

# Paired-End mode is off

INFO @ Sun, 13 Mar 2022 19:18:00: #1 read tag files...

INFO @ Sun, 13 Mar 2022 19:18:00: #1 read treatment tags...

INFO @ Sun, 13 Mar 2022 19:18:10: 10000000

INFO @ Sun, 13 Mar 2022 19:18:13: 13000000

INFO @ Sun, 13 Mar 2022 19:18:17: 18000000

INFO @ Sun, 13 Mar 2022 19:18:29: 30000000

INFO @ Sun, 13 Mar 2022 19:18:30: 31000000

INFO @ Sun, 13 Mar 2022 19:18:37: #1.2 read input tags...

INFO @ Sun, 13 Mar 2022 19:18:40: 3000000

INFO @ Sun, 13 Mar 2022 19:18:44: 7000000

INFO @ Sun, 13 Mar 2022 19:19:03: 23000000

INFO @ Sun, 13 Mar 2022 19:19:07: 26000000

INFO @ Sun, 13 Mar 2022 19:19:29: 45000000

INFO @ Sun, 13 Mar 2022 19:19:33: #1 tag size is determined as 76 bps

INFO @ Sun, 13 Mar 2022 19:19:33: #1 tag size = 76.0

INFO @ Sun, 13 Mar 2022 19:19:33: #1 total tags in treatment: 4098003

INFO @ Sun, 13 Mar 2022 19:19:33: #1 user defined the maximum tags...

INFO @ Sun, 13 Mar 2022 19:19:33: #1 filter out redundant tags at the same location and the same strand by allowing at most 1 tag(s)

INFO @ Sun, 13 Mar 2022 19:19:33: #1 tags after filtering in treatment: 3324960

INFO @ Sun, 13 Mar 2022 19:19:33: #1 Redundant rate of treatment: 0.19

INFO @ Sun, 13 Mar 2022 19:19:33: #1 total tags in control: 6242924

INFO @ Sun, 13 Mar 2022 19:19:33: #1 user defined the maximum tags...

INFO @ Sun, 13 Mar 2022 19:19:33: #1 filter out redundant tags at the same location and the same strand by allowing at most 1 tag(s)

INFO @ Sun, 13 Mar 2022 19:19:33: #1 tags after filtering in control: 4676276

INFO @ Sun, 13 Mar 2022 19:19:33: #1 Redundant rate of control: 0.25

INFO @ Sun, 13 Mar 2022 19:19:33: #1 finished!

INFO @ Sun, 13 Mar 2022 19:19:33: #2 Build Peak Model...

INFO @ Sun, 13 Mar 2022 19:19:33: #2 looking for paired plus/minus strand peaks...

INFO @ Sun, 13 Mar 2022 19:19:35: #2 number of paired peaks: 40704

INFO @ Sun, 13 Mar 2022 19:19:35: start model\_add\_line...

INFO @ Sun, 13 Mar 2022 19:19:35: start X-correlation...

INFO @ Sun, 13 Mar 2022 19:19:35: end of X-cor

INFO @ Sun, 13 Mar 2022 19:19:35: #2 finished!

INFO @ Sun, 13 Mar 2022 19:19:35: #2 predicted fragment length is 75 bps

INFO @ Sun, 13 Mar 2022 19:19:35: #2 alternative fragment length(s) may be 75 bps

INFO @ Sun, 13 Mar 2022 19:19:35: #2.2 Generate R script for model : macs2/NA\_model.r

WARNING @ Sun, 13 Mar 2022 19:19:35: #2 Since the d (75) calculated from paired-peaks are smaller than 2\*tag length, it may be influenced by unknown sequencing problem!

WARNING @ Sun, 13 Mar 2022 19:19:35: #2 You may need to consider one of the other alternative d(s): 75

WARNING @ Sun, 13 Mar 2022 19:19:35: #2 You can restart the process with --nomodel --extsize XXX with your choice or an arbitrary number. Nontheless, MACS will continute computing.

INFO @ Sun, 13 Mar 2022 19:19:35: #3 Call peaks...

INFO @ Sun, 13 Mar 2022 19:19:35: #3 Call broad peaks with given level1 -log10qvalue cutoff and level2: 1.301030, 1.000000...

INFO @ Sun, 13 Mar 2022 19:19:35: #3 Pre-compute pvalue-qvalue table...

INFO @ Sun, 13 Mar 2022 19:19:59: #3 Call peaks for each chromosome...

INFO @ Sun, 13 Mar 2022 19:20:13: #4 Write output xls file... macs2/NA\_peaks.xls

INFO @ Sun, 13 Mar 2022 19:20:13: #4 Write broad peak in broadPeak format file... macs2/NA\_peaks.broadPeak

INFO @ Sun, 13 Mar 2022 19:20:13: #4 Write broad peak in bed12/gappedPeak format file... macs2/NA\_peaks.gappedPeak

INFO @ Sun, 13 Mar 2022 19:20:13: Done!

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ gunzip ENCFF806LPY.bed.gz

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ intervene venn -i macs2/AFY\_WIS/NA\_peaks\_ENCFF182AFY.broadPeak ENCFF806LPY.bed --filenames --output venn\_results/AFY\_WIS/venn1.jpg

Generating a 2-way "venn" diagram. Please wait...

Done! Please check your results @ venn\_results/AFY\_WIS/venn1.jpg.

Thank you for using Intervene!

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ intervene venn -i ENCFF806LPY.bed macs2/AFY\_WIS/NA\_peaks\_ENCFF182AFY.broadPeak --filenames --output venn\_results/AFY\_WIS/venn2.jpg

Generating a 2-way "venn" diagram. Please wait...

Done! Please check your results @ venn\_results/AFY\_WIS/venn2.jpg.

Thank you for using Intervene!

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ intervene venn -i macs2/CQJ\_WIS/NA\_peaks\_ENCFF335CQJ.broadPeak ENCFF806LPY.bed --filenames --output venn\_results/CQJ\_WIS/venn1.jpg

Generating a 2-way "venn" diagram. Please wait...

Done! Please check your results @ venn\_results/CQJ\_WIS/venn1.jpg.

Thank you for using Intervene!

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$ intervene venn -i ENCFF806LPY.bed macs2/CQJ\_WIS/NA\_peaks\_ENCFF335CQJ.broadPeak --filenames --output venn\_results/CQJ\_WIS/venn2.jpg

Generating a 2-way "venn" diagram. Please wait...

Done! Please check your results @ venn\_results/CQJ\_WIS/venn2.jpg.

Thank you for using Intervene!

(base) leonid@leonid-Aspire-A315-42:~/HW2\_sem4$