#clear the generated data

rm -fvr test\* cLoops2.log gm\* k562\*

#1. get basic statistics of PETs

cLoops2 qc -f ../data/GM\_HiTrac\_bio1.bedpe.gz,../data/GM\_HiTrac\_bio2.bedpe.gz,../data/K562\_HiTrac\_bio1.bedpe.gz,../data/K562\_HiTrac\_bio2.bedpe.gz -o test\_step1 -p 4

#2. pre-process BEDPE file to cLoops2 data

##get directory seperately for GM12878, only target chromosome chr21

cLoops2 pre -f ../data/GM\_HiTrac\_bio1.bedpe.gz -o gm\_bio1 -c chr21

cLoops2 pre -f ../data/GM\_HiTrac\_bio2.bedpe.gz -o gm\_bio2 -c chr21

#get the combined data for GM12878

cLoops2 pre -f ../data/GM\_HiTrac\_bio1.bedpe.gz,../data/GM\_HiTrac\_bio2.bedpe.gz -o gm -c chr21

##get the directory seperately for K562 first

cLoops2 pre -f ../data/K562\_HiTrac\_bio1.bedpe.gz -o k562\_bio1 -c chr21

cLoops2 pre -f ../data/K562\_HiTrac\_bio2.bedpe.gz -o k562\_bio2 -c chr21

##then combine the data, only keep 1 PET for the same position, default the same to cLoops2 pre

cLoops2 combine -ds k562\_bio1,k562\_bio2 -o k562 -keep 1

#3. estimate reasonable resolution

cLoops2 estRes -d gm -o gm -bs 5000,1000,200 -p 10

cLoops2 estRes -d k562 -o k562 -bs 5000,1000,200 -p 10

#4. estimate interaction distance

cLoops2 estDis -d gm -o gm -bs 1000 -p 10 -plot

#5. estimate data similarities

cLoops2 estSim -ds gm\_bio1,gm\_bio2,gm,k562\_bio1,k562\_bio2,k562 -bs 1000 -plot -p 6 -o test\_step4

#6. call peaks

#can call directly, or run cLoops2 filterPETs filtering first, will better clean/fast

cLoops2 callPeaks -d gm -o gm -eps 50,100 -minPts 10 -mcut 1000 -split

#7. show aggreated peaks

cLoops2 agg -d gm -peaks gm\_peaks.bed -o gm -peak\_ext 2500 -peak\_bins 200 -peak\_norm -skipZeros

#7.1 show aggregated view points

cLoops2 agg -d gm -viewPoints gm\_peaks.bed -o gm -viewPointBs 1000 -viewPointDown 30000 -viewPointUp 20000 -1D -bws ../data/GM12878\_CTCF\_chr21.bw

#8. call intra-chromosomal loops, filtered PETs can be used to show clear view of loops, or futhur to call loops

cLoops2 callLoops -d gm -o gm -eps 200,500,1000 -minPts 10 -w -j

#9. show aggreated loops, with -bws option ATAC-seq/ChIP-seq tracks can also be shown

cLoops2 agg -d gm -o gm -loops gm\_loops.txt -bws ../data/GM12878\_ATAC\_chr21.bw,../data/GM12878\_CTCF\_chr21.bw -1D -loop\_norm

#9.1 show two anchors

cLoops2 agg -d gm -o gm -twoAnchors gm\_loops.txt -1D -bws ../data/GM12878\_CTCF\_chr21.bw,../data/GM12878\_ATAC\_chr21.bw -twoAnchor\_ext 0.5

#10. call domains

cLoops2 callDomains -d gm -o gm -bs 5000 -ws 100000,250000

#11. show aggregated domains

#convert the output segregation score from bedGraph file to bigWig

bedGraphToBigWig gm\_domains\_SS\_binSize5.0k\_winSize100.0k.bdg ../../data/hg38.chrom.sizes gm\_domains\_SS\_bs5k\_ws100k.bw

bedGraphToBigWig gm\_domains\_SS\_binSize5.0k\_winSize250.0k.bdg ../../data/hg38.chrom.sizes gm\_domains\_SS\_bs5k\_ws250k.bw

cLoops2 agg -d gm -o gm -domains gm\_domains.bed -bws ../data/GM12878\_CTCF\_chr21.bw,gm\_domains\_SS\_bs5k\_ws100k.bw,gm\_domains\_SS\_bs5k\_ws250k.bw -1D

#12. visualization

#show example interactions

cLoops2 plot -f ./gm/chr21-chr21.ixy -o gm\_domain\_example -bs 5000 -start 35800000 -end 36700000 -domains gm\_domains.bed -log -bws ../data/GM12878\_CTCF\_chr21.bw -1D -corr

#show enhancer-promoter loops

cLoops2 plot -f gm/chr21-chr21.ixy -o gm\_example -bs 500 -start 38752604 -end 38839334 -triu -bws ../data/GM12878\_ATAC\_chr21.bw,../data/GM12878\_CTCF\_chr21.bw -1D -loops gm\_loops.txt -beds ../data/GM12878\_RoadMap\_hg38\_Enh\_chr21.bed,../data/GM12878\_RoadMap\_hg38\_Tss\_chr21.bed,gm\_peaks.bed -m obs -log -gtf ../data/gencode\_v30\_chr21.gtf -vmax 1

cLoops2 plot -o gm\_example -bs 500 -chrom chr21 -start 38752604 -end 38839334 -bws ../data/GM12878\_ATAC\_chr21.bw,../data/GM12878\_CTCF\_chr21.bw -1D -loops gm\_loops.txt -beds ../data/GM12878\_RoadMap\_hg38\_Enh\_chr21.bed,../data/GM12878\_RoadMap\_hg38\_Tss\_chr21.bed,gm\_peaks.bed -gtf ../data/gencode\_v30\_chr21.gtf

#filter data

cLoops2 filterPETs -d gm -loops gm\_loops.txt -o gm\_filtered

cLoops2 plot -f gm\_filtered/chr21-chr21.ixy -o gm\_filtered\_example -bs 500 -start 38752604 -end 38839334 -triu -loops gm\_loops.txt -log -1D

#more clear of arches

cLoops2 plot -f gm\_filtered/chr21-chr21.ixy -o gm\_example -start 46228500 -end 46290000 -1D -loops gm\_loops.txt -arch -aw 0.5

#13. differentially enriched loops

#a. sampling PETs to same/similar depth to call loops with same parameters

cLoops2 samplePETs -d gm -o gm\_samp -tot 780000

cLoops2 samplePETs -d k562 -o k562\_samp -tot 780000

#b. call loops with same parameters

cLoops2 callLoops -d gm\_samp -o gm\_samp -eps 200,500,1000 -minPts 10 -w -j

cLoops2 callLoops -d k562\_samp -o k562\_samp -eps 200,500,1000 -minPts 10 -w -j

#c. call differentially enriched loops

cLoops2 callDiffLoops -tloop gm\_samp\_loops.txt -td gm\_samp -cloop k562\_samp\_loops.txt -cd k562\_samp -o gm\_vs\_k562 -j -w

#14. cLoops2 data to other

#convert data to bed file, used for other tools

cLoops2 dump -d gm -o gm -bed

#convert data to bedpe file, used for other tools

cLoops2 dump -d gm -o gm -bedpe

#convert data to 1D track, can be furthur convert to bigWig file

cLoops2 dump -d gm -o gm -bdg

#convert data to washU track

cLoops2 dump -d gm -o gm -washU

#convert data to .hic file

cLoops2 dump -d gm -o gm -hic -hic\_org hg38 -hic\_res 200000,25000,5000

#convert data to matrix

cLoops2 dump -d gm -mat -o gm -mat\_res 10000 -mat\_chrom chr21-chr21 -mat\_start 36000000 -mat\_end 40000000 -log -norm -corr

#15. quantify feautres

#quantify GM12878 peaks in K562 data

cLoops2 quant -d k562 -peaks gm\_peaks.bed -o k562\_gm

#quantify GM12878 loops in K562 data

cLoops2 quant -d k562 -loops gm\_loops.txt -o k562\_gm

#quantify GM12878 domains in K562 data

cLoops2 quant -d k562 -domains gm\_domains.txt -o k562\_gm -domain\_bs 5000 -domain\_ws 100000

#16. montage analysis

#all interactions

cLoops2 montage -f ./gm\_samp/chr21-chr21.ixy -bed ../data/runx1.bed -o gm\_runx1\_all -ext 2 -simple -ppmw 0.05 -vmax 500

cLoops2 montage -f ./k562\_samp/chr21-chr21.ixy -bed ../data/runx1.bed -o k562\_runx1\_all -ext 2 -simple -ppmw 0.05 -vmax 500

#viewpoints mode for E19 and promoter

cLoops2 montage -f ./gm\_samp/chr21-chr21.ixy -bed ../data/runx1.bed -o gm\_runx1\_vp -ext 2 -simple -ppmw 0.05 -vmax 500 -vp Promoter,E19

cLoops2 montage -f ./k562\_samp/chr21-chr21.ixy -bed ../data/runx1.bed -o k562\_runx1\_vp -ext 2 -simple -ppmw 0.05 -vmax 500 -vp Promoter,E19

#17. anotating loops

cLoops2 anaLoops -loop gm\_loops.txt -o gm\_loops -gtf ../data/gencode\_v30\_chr21.gtf -net

#show networks

python plotNetExample.py

############analysis related functions##########

#get the singal distribution

getSigDist.py -d gm -o gm -r 5 -plot -log

#get the segregation score

getSS.py -f gm/chr21-chr21.ixy -o gm\_chr21