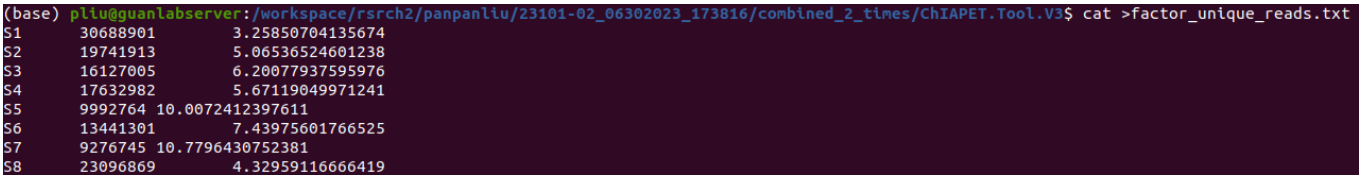
normalized loops number by mapped reads

factor= 1*1e8/unique_reads_number

normailized by unique mapped reads number for pet detection(1*1e8/u

S1	30688901	3.25850704135674
S2	19741913	5.06536524601238
S3	16127005	6.20077937595976
S4	17632982	5.67119049971241
S5	9992764	10.0072412397611
S6	13441301	7.43975601766525
S7	9276745	10.7796430752381
S8	23096869	4.32959116666419

cd /workspace/rsrch2/panpanliu/23101-02_06302023_173816/combined_2_times/ChIAPET.Tool.V3
cat >factor_unique_reads.txt
S1 30688901 3.25850704135674
S2 19741913 5.06536524601238
S3 16127005 6.20077937595976
S4 17632982 5.67119049971241
S5 9992764 10.0072412397611
S6 13441301 7.43975601766525
S7 9276745 10.7796430752381
S8 23096869 4.32959116666419



get the normalized pet counts number

```
while read -r line
 do
 sam=$(echo $line | awk '{print $1}')
     factor=$(echo $line | awk '{print $3}')
     echo -e "$sam\t$factor"
cd /workspace/rsrch2/panpanliu/23101-02_06302023_173816/combined_2_times/ChIAPET.Tool.V3/${sam}
awk '{print $1"\t"$2"\t"$3"\t"$4"\t"$5"\t"$6"\t"($7*'$factor')}' ${sam}.cluster.FDRfiltered.txt >./normalizaed_by_uniqreads/${sam}.uniqreads_normlized.cluster.FDRfiltered.txt
done <factor_unique_reads.txt
```

normalizated by mapped reads

cd /workspace/rsrch2/panpanliu/23101-02_06302023_173816/combined_2_times/ChIAPET.Tool.V3
factor_mapped_reads.txt

```
while read -r line
 do
 sam=$(echo $line | awk '{print $1}')
     factor=$(echo $line | awk '{print $3}')
     echo -e "$sam\t$factor"
cd /workspace/rsrch2/panpanliu/23101-02_06302023_173816/combined_2_times/ChIAPET.Tool.V3/${sam}
mkdir normalized_mapped_reads
awk '{print $1"\t"$2"\t"$3"\t"$4"\t"$5"\t"$6"\t"($7*'$factor')}' ${sam}.cluster.FDRfiltered.txt >./normalized_mapped_reads/${sam}.mapReads_normlized.cluster.FDRfiltered.txt
done <factor_mapped_reads.txt
```

loops num across circadian genes###

```
for f in $(ls -d S[1-8]/*.cluster.FDRfiltered.txt)
 do
 echo $f

 sam=`basename $f .cluster.FDRfiltered.txt`

 pairToBed -a /workspace/rsrch2/panpanliu/23101-02_06302023_173816/combined_2_times/ChIAPET.Tool.V3/${sam}/normalized_mapped_reads/${sam}.mapReads_normlized.cluster.FDRfiltered.txt -b /workspace/rsrch2/common_data/Refgenome/mm10/homer_transcript_fetchsize.TSS2.5k.sorted.bed >/workspace/rsrch2/panpanliu/23101-02_06302023_173816/combined_2_times/ChIAPET.Tool.V3/${sam}/normalized_mapped_reads/${sam}_norm_transTSS2.5kb_interscet.txt
 done
```

step 4: use R script to group by gene/transcript to count the number cross transcripts
cd /workspace/rsrch2/panpanliu/23101-02_06302023_173816/combined_2_times/ChIAPET.Tool.V3
merge_file_norm_loop.R

```
rm(list = ls())
setwd("/workspace/rsrch2/panpanliu/23101-02_06302023_173816/combined_2_times/ChIAPET.Tool.V3")
library(dplyr)
library(data.table)

cutoff="2.5" ## change this number to plot different plots with diffrent
```

```
values#

pre="all_transcrpts_norm"

B6NCZT22_loop <-
read.delim("S2/normalized_mapped_reads/S2_norm_transTSS2.5kb_interscet.txt",str
ingsAsFactors = F,header = F) %>%
  filter(V11 != "0" & V7 >= cutoff)
B6NCZT22_counts <- B6NCZT22_loop %>%
  group_by(V11,V12) %>%
  summarise(total_count=n(),.groups = 'drop') %>%
  as.data.frame()
rownames(B6NCZT22_counts)=paste0(B6NCZT22_counts$V12,"|",B6NCZT22_counts$V11)
colnames(B6NCZT22_counts)=c("gene_symbol","transcriptID","B6NCZT22")

S129NCZT22_loop <-
read.delim("S6/normalized_mapped_reads/S6_norm_transTSS2.5kb_interscet.txt",str
ingsAsFactors = F,header = F) %>%
  filter(V11 != "0" & V7 >= cutoff)
S129NCZT22_counts <- S129NCZT22_loop %>%
  group_by(V11,V12) %>%
  summarise(total_count=n(),.groups = 'drop') %>%
  as.data.frame()
rownames(S129NCZT22_counts)=paste0(S129NCZT22_counts$V12,"|",S129NCZT22_counts$
V11)
colnames(S129NCZT22_counts)=c("gene_symbol","transcriptID","S129NCZT22")




##fgene_list
#prefix="common_cir"
#common_cir<-
read.csv("/workspace/rsrch2/common_data/B6_129_starin_res/Circadian_genes_snp/N
C_Common_Cir.csv")

prefix="S129_spe"
common_cir<-
read.csv("/workspace/rsrch2/common_data/B6_129_starin_res/Circadian_genes_snp/S
129_Specific_Cir_FC2_P0.05.csv")
rownames(common_cir)=common_cir$CycID
temp <- merge(common_cir,B6NCZT22_counts,by = 'row.names',all.x = T)
temp <- subset(temp,select=-c(1,(dim(temp)[2]-2),(dim(temp)[2]-1)))
rownames(temp)=temp$CycID

common_cir_loop <- merge(temp, S129NCZT22_counts,by='row.names',all.x=T)
common_cir_loop <-  subset(common_cir_loop,select=-c(1,(dim(common_cir_loop)
[2]-2),(dim(common_cir_loop)[2]-1)))




#common_gene<-
intersect(common_cir$CycID,intersect(rownames(B6NCZT22_counts),rownames(S129NCZ
T22_counts)))
```

```
#common_cir_loop <-
cbind(common_cir[common_gene,],B6NCZT22_counts[common_gene,"B6NCZT22"],S129NCZT
22_counts[common_gene,"S129NCZT22"])
colnames(common_cir_loop)=c(colnames(common_cir),"B6NCZT22_loops","S129NCZT22_l
oops")

write.table(common_cir_loop,file=paste0("/workspace/rsrch2/common_data/B6_129_s
tarin_res/norlized_loopsbymappedreads/",prefix,"loops_num.txt"),
            sep = "\t",quote = F,row.names = F,na="0")


boxplotdata<-reshape2::melt(common_cir_loop[,c(1,10,11)],id="CycID",value.name
= "loops_num")
class(boxplotdata$variable)
library(ggplot2)
ggplot(boxplotdata, aes(x=variable, y=loops_num,color=variable)) +
  geom_boxplot(outlier.colour="black", outlier.shape=16,
               outlier.size=1, notch=FALSE)+
  labs(title=paste0("Plot of loops for ",pre),x="sample", y = "Loops_num")+
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))




##









###############not run###############3



##
```

```
NC_B6<-merge(B6NCZT10,B6NCZT22,by="V1",all=T)
colnames(NC_B6)=c("gene","B6NCZT10","B6NCZT22")
HF_B6<-merge(B6HFZT10,B6HFZT10,by="V1",all=T)
colnames(HF_B6)=c("gene","B6HFZT10","B6HFZT10")
B6 <- merge(NC_B6,HF_B6,by="gene",all=T)

NC_S129 <- merge(S129NCZT10,S129NCZT22,by="V1",all=T)
colnames(NC_S129)=c("gene","S129NCZT10","S129NCZT22")
HF_S129 <- merge(S129HFZT10,S129HFZT22,by="V1",all=T)
colnames(HF_S129)=c("gene","S129HFZT10","S129HFZT22")
S129 <- merge(NC_S129,HF_S129,by="gene",all=T)

B6_S129 <- merge (B6,S129, by="gene",all=T)
rownames(B6_S129)=B6_S129$gene

write.table(B6_S129,file=paste0(pre,"_intersect_loops_num.txt"),quote=F,sep =
"\t",
            na = "0",row.names = F)



#####plot log2FC (gene_Expression) and loops number correlation









##NC ZT22
DEG<-read.csv("NC_B6_129_ZT22_DEG.csv")
colnames(DEG)[1]=c("gene")
DEG$FC=2^DEG$log2FC
rownames(DEG)=DEG$gene

genestring<-
intersect(DEG$gene,intersect(rownames(B6NCZT22_counts),rownames(S129NCZT22_coun
```

```
ts)))

plotdata <-
cbind(DEG[genestring,c("gene","FC")],B6NCZT22_counts[genestring,"B6NCZT22"],S12
9NCZT22_counts[genestring,"S129NCZT22"])
plotdata <- na.omit(plotdata)


plotdata$loopsFC<-plotdata$S129NCZT22/plotdata$B6NCZT22

library(ggplot2)
ggplot(plotdata, aes(x = FC, y = loopsFC)) +
  geom_point()+
  scale_x_continuous(limits = c(0, 20))+
  scale_y_continuous(limits = c(0, 20))+
  ggtitle(paste0("normcounts for DEGs at",cutoff))

cor(plotdata$FC,plotdata$loopsFC)









boxplotdata<-reshape2::melt(B6_S129,id="gene",value.name = "loops_num")
class(boxplotdata$variable)

ggplot(boxplotdata, aes(x=variable, y=loops_num,color=variable)) +
  geom_boxplot(outlier.colour="black", outlier.shape=16,
               outlier.size=1, notch=FALSE)+
  labs(title=paste0("Plot of loops for ",pre),x="sample", y = "Loops_num")+
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```