Code to generate Figure 5 of 'Glucocorticoid receptor collaborates with pioneer factors and AP-1 to execute genome-wide regulation'

Erin M. Wissink

```
library(ggplot2)
library(bigWig)
library(gridExtra)
library(UpSetR)
library(bedtoolsr)

ssBigwigPath <- '../Manuscript_data/PRO-seq_bigwigs/Merged_subsampled/'
msDataPath <- '../Manuscript_data/'


geneData <- read.table('gene_DESeq_analysis.txt', sep = '\t')
promoters <- read.table(paste0(msDataPath, 'hg38_refseq_promoters.bed'), sep = '\t')
promA549 <-
  bedtoolsr::bt.sort(i = read.table('Promoter_highest_read_count_A549.bed', sep = '\t'))
promU2OS <-
  bedtoolsr::bt.sort(i = read.table('Promoter_highest_read_count_U2OS.bed', sep = '\t'))
gg_options <- theme(axis.text.x = element_text(size = 8),
                    axis.text.y = element_text(size = 8),
                    axis.title.x = element_text(size = 8, colour='black'),
                    axis.title.y = element_text(size = 8, colour='black'),
                    legend.text=element_text(size=8, colour='black'),
                    legend.title = element_text(size = 8, colour='black'),
                    axis.line = element_line(colour = 'black', size = 0.5),
                    axis.ticks = element_line(colour = "black", size = 0.5)) +
                    theme_classic()
```

Looking at overlap of GR binding and dex-responsive CCREs

```
CCREs <- read.table('CCRE_DESeq_analysis.txt')[,c(14:17, 19:22)]

#for CCREs that were not tested because of low read counts, set padj to 1
CCREs$A549_wt_1vs0dex_fdr[is.na(CCREs$A549_wt_1vs0dex_fdr)] <- 1
CCREs$A549_wt_100vs0dex_fdr[is.na(CCREs$A549_wt_100vs0dex_fdr)] <- 1
CCREs$U2OS_wt_1vs0dex_fdr[is.na(CCREs$U2OS_wt_1vs0dex_fdr)] <- 1
CCREs$U2OS_wt_100vs0dex_fdr[is.na(CCREs$U2OS_wt_100vs0dex_fdr)] <- 1

#for CCREs that were not tested because of low read counts, set log2fc to 0
CCREs[is.na(CCREs)] <- 0
CCREs$category <- ''
```

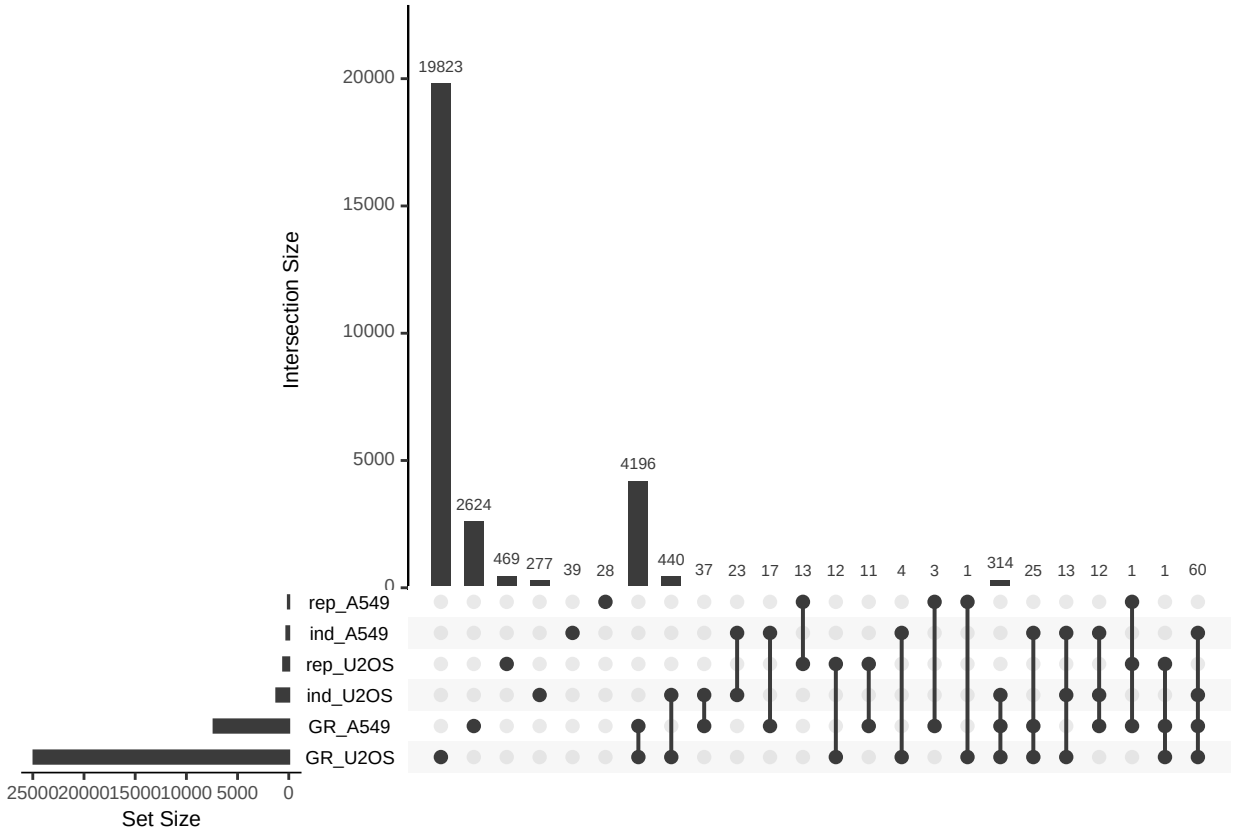
```
for (i in 1:nrow(CCREs)){
  if (CCREs[i,4] < 0.05 & CCREs[i,8]<0.05){
    if (CCREs[i,3] > 0 & CCREs[i,7] > 0){
      CCREs[i,9] <- 'both_up'
    }
    else if (CCREs[i,3] < 0 & CCREs[i,7] < 0){
      CCREs[i,9] <- 'both_down'
    }
    else{
      CCREs[i,9] <- 'opposite'
    }
  }
  else if (CCREs[i,4] <0.05){
    if (CCREs[i,3] > 0) {
      CCREs[i,9] <- 'A549_up'
    }
    else {CCREs[i,9] <- 'A549_down'}
  }
  else if (CCREs[i,8] < 0.05){
    if (CCREs[i,7] > 0){
      CCREs[i,9] <- 'U2OS_up'
    }
    else{CCREs[i,9] <- 'U2OS_down'}
  }
  else{CCREs[i,9] <- 'bkgd'}
}

CCREs$cood <- rownames(CCREs)
CCREs$chr <- ''
CCREs$start <- ''
CCREs$end <- ''

for (i in 1:nrow(CCREs)){
  CCREs[i,11] <- substr(CCREs[i,10], 1,
                        which(strsplit(CCREs[i,10], "")[[1]]==":")-1)
  CCREs[i,12] <- substr(CCREs[i,10],
                        which(strsplit(CCREs[i,10], "")[[1]]==":")+1,
                        which(strsplit(CCREs[i,10], "")[[1]]=="-")-1)
  CCREs[i,13] <- substr(CCREs[i,10],
                        which(strsplit(CCREs[i,10], '')[[1]]=='-')+1,
                        stop = nchar(CCREs[i,10]))
}

CCREs$start500 <- (as.numeric(CCREs$start) + as.numeric(CCREs$end))/2 -250
CCREs$end500 <- CCREs$start500 + 500

A549_up_all_bed <- CCREs[CCREs$category=='both_up' |
                           CCREs$category=='A549_up', c(11, 14, 15)]
U2OS_up_all_bed <- CCREs[CCREs$category=='both_up' |
                           CCREs$category=='U2OS_up', c(11, 14, 15)]
A549_down_all_bed <- CCREs[CCREs$category=='both_down' |
                           CCREs$category=='A549_down', c(11, 14, 15)]
U2OS_down_all_bed <- CCREs[CCREs$category=='both_down' |
```

```
                                CCREs$category=='U2OS_down', c(11, 14, 15)]
```

```
#Reading in data for GR binding sites

A549_U2OS_GR_overlap <- read.table('A549_U2OS_GR_overlap.bed', sep = '\t')
A549_U2OS_GR_overlap$GR_A549 <- 1
A549_U2OS_GR_overlap$GR_U2OS <- 1
A549_U2OS_GR_overlap$ind_A549 <- 0
A549_U2OS_GR_overlap$ind_U2OS <- 0
A549_U2OS_GR_overlap$rep_A549 <- 0
A549_U2OS_GR_overlap$rep_U2OS <- 0
A549_GR <- read.table('A549_GR_unique.bed', sep = '\t')
A549_GR$GR_A549 <- 1
A549_GR$GR_U2OS <- 0
A549_GR$ind_A549 <- 0
A549_GR$ind_U2OS <- 0
A549_GR$rep_A549 <- 0
A549_GR$rep_U2OS <- 0
U2OS_GR <- read.table('U2OS_GR_unique.bed', sep = '\t')
U2OS_GR$GR_A549 <- 0
U2OS_GR$GR_U2OS <- 1
U2OS_GR$ind_A549 <- 0
U2OS_GR$ind_U2OS <- 0
U2OS_GR$rep_A549 <- 0
U2OS_GR$rep_U2OS <- 0

toUpSet <- rbind(A549_U2OS_GR_overlap, A549_GR, U2OS_GR)

#I want to expand the binding sites from 50 bp to 500 bp
# for overlapping with CCREs
toUpSet$V2 <- toUpSet$V2 - 225
toUpSet$V3 <- toUpSet$V3 + 224

rownames(toUpSet) <-
  paste(toUpSet$V1, ':', toUpSet$V2, '-', toUpSet$V3)


A549_ind_GR_overlap <- bedtoolsr::bt.intersect(wa = T, a = toUpSet, b = A549_up_all_bed)
A549_ind_GR_overlap <-
  paste(A549_ind_GR_overlap$V1, ':', A549_ind_GR_overlap$V2, '-', A549_ind_GR_overlap$V3)
toUpSet[A549_ind_GR_overlap, 6] <- 1

A549_ind_no_GR <- bedtoolsr::bt.intersect(v = T, a = A549_up_all_bed, b = toUpSet)
rownames(A549_ind_no_GR) <-
  paste(A549_ind_no_GR$V1, ':', A549_ind_no_GR$V2, '-', A549_ind_no_GR$V3)
A549_ind_no_GR$GR_A549 <- 0
A549_ind_no_GR$GR_U2OS <- 0
A549_ind_no_GR$ind_A549 <- 1
A549_ind_no_GR$ind_U2OS <- 0
A549_ind_no_GR$rep_A549 <- 0
A549_ind_no_GR$rep_U2OS <- 0
toUpSet <- rbind(toUpSet, A549_ind_no_GR)
```

```
A549_rep_GR_overlap <- bedtoolsr::bt.intersect(wa = T, a = toUpSet, b = A549_down_all_bed)
A549_rep_GR_overlap <-
  paste(A549_rep_GR_overlap$V1, ':', A549_rep_GR_overlap$V2, '-', A549_rep_GR_overlap$V3)
toUpSet[A549_rep_GR_overlap, 8] <- 1

A549_rep_no_GR <- bedtoolsr::bt.intersect(v = T, a = A549_down_all_bed, b = toUpSet)
rownames(A549_rep_no_GR) <-
  paste(A549_rep_no_GR$V1, ':', A549_rep_no_GR$V2, '-', A549_rep_no_GR$V3)
A549_rep_no_GR$GR_A549 <- 0
A549_rep_no_GR$GR_U2OS <- 0
A549_rep_no_GR$ind_A549 <- 0
A549_rep_no_GR$ind_U2OS <- 0
A549_rep_no_GR$rep_A549 <- 1
A549_rep_no_GR$rep_U2OS <- 0
toUpSet <- rbind(toUpSet, A549_rep_no_GR)

U2OS_ind_GR_overlap <- bedtoolsr::bt.intersect(wa = T, a = toUpSet, b = U2OS_up_all_bed)
U2OS_ind_GR_overlap <-
  paste(U2OS_ind_GR_overlap$V1, ':', U2OS_ind_GR_overlap$V2, '-', U2OS_ind_GR_overlap$V3)
toUpSet[U2OS_ind_GR_overlap, 7] <- 1

U2OS_ind_no_GR <- bedtoolsr::bt.intersect(v = T, a = U2OS_up_all_bed, b = toUpSet)
rownames(U2OS_ind_no_GR) <-
  paste(U2OS_ind_no_GR$V1, ':', U2OS_ind_no_GR$V2, '-', U2OS_ind_no_GR$V3)
U2OS_ind_no_GR$GR_A549 <- 0
U2OS_ind_no_GR$GR_U2OS <- 0
U2OS_ind_no_GR$ind_A549 <- 0
U2OS_ind_no_GR$ind_U2OS <- 1
U2OS_ind_no_GR$rep_A549 <- 0
U2OS_ind_no_GR$rep_U2OS <- 0
toUpSet <- rbind(toUpSet, U2OS_ind_no_GR)

U2OS_rep_GR_overlap <- bedtoolsr::bt.intersect(wa = T, a = toUpSet, b = U2OS_down_all_bed)
U2OS_rep_GR_overlap <-
  paste(U2OS_rep_GR_overlap$V1, ':', U2OS_rep_GR_overlap$V2, '-', U2OS_rep_GR_overlap$V3)
toUpSet[U2OS_rep_GR_overlap, 9] <- 1

U2OS_rep_no_GR <- bedtoolsr::bt.intersect(v = T, a = U2OS_down_all_bed, b = toUpSet)
rownames(U2OS_rep_no_GR) <-
  paste(U2OS_rep_no_GR$V1, ':', U2OS_rep_no_GR$V2, '-', U2OS_rep_no_GR$V3)
U2OS_rep_no_GR$GR_A549 <- 0
U2OS_rep_no_GR$GR_U2OS <- 0
U2OS_rep_no_GR$ind_A549 <- 0
U2OS_rep_no_GR$ind_U2OS <- 0
U2OS_rep_no_GR$rep_A549 <- 0
U2OS_rep_no_GR$rep_U2OS <- 1
toUpSet <- rbind(toUpSet, U2OS_rep_no_GR)

chrom <- c('chr1','chr2','chr3','chr4','chr5', 'chr6', 'chr7', 'chr8', 'chr9',
           'chr10', 'chr11', 'chr12', 'chr13', 'chr14', 'chr15', 'chr16', 'chr17',
           'chr18', 'chr19', 'chr20', 'chr21', 'chr22', 'chrX')
toUpSet <- toUpSet[toUpSet$V1 %in% chrom,]
```

```
upset(toUpSet, nsets = 6)
```




```
table(toUpSet$GR_A549)
```

```
##
##     0     1
## 21142  7301
```

```
table(toUpSet$GR_U2OS)
```

```
##
##     0     1
##  3554 24889
```

```
table(toUpSet$ind_A549)
```

```
##
##     0     1
## 28250   193
```

```
table(toUpSet$ind_U2OS)
```

```
##
##     0     1
## 27267  1176
```

```
table(toUpSet$rep_A549)
```

```
##
##     0     1
## 28397    46
```

```
table(toUpSet$rep_U2OS)
```

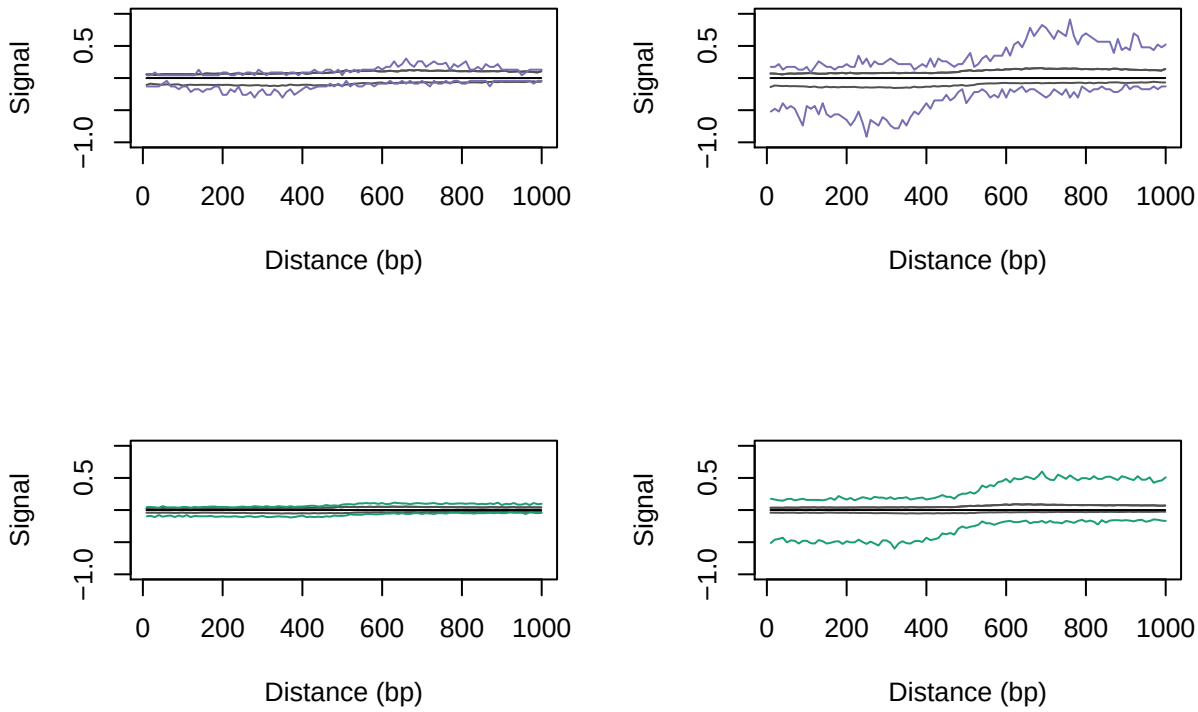
```
##
##     0     1
## 27936   507
```

PRO-seq metaplots at GORS with and without induced CCREs

```
# Making BED files from above, extending to +-500 bp, and adding both strands
A549_GR_A549_induced <- toUpSet[toUpSet$GR_A549 == 1 & toUpSet$ind_A549 == 1, 1:3]
A549_GR_A549_induced <- center.bed(A549_GR_A549_induced, 500,499)
A549_GR_A549_induced_bed <- rbind(A549_GR_A549_induced, A549_GR_A549_induced)
A549_GR_A549_induced_bed$V4 <- '.'
A549_GR_A549_induced_bed$V5 <- '.'
A549_GR_A549_induced_bed$V6 <-
  c(rep('+', nrow(A549_GR_A549_induced)), rep('-', nrow(A549_GR_A549_induced)))

A549_GR_no_changed_CCREs <- toUpSet[toUpSet$GR_A549 == 1 & toUpSet$ind_A549 == 0, 1:3]
A549_GR_no_changed_CCREs <- center.bed(A549_GR_no_changed_CCREs, 500,499)
A549_GR_no_changed_bed <- rbind(A549_GR_no_changed_CCREs, A549_GR_no_changed_CCREs)
A549_GR_no_changed_bed$V4 <- '.'
A549_GR_no_changed_bed$V5 <- '.'
A549_GR_no_changed_bed$V6 <-
  c(rep('+', nrow(A549_GR_no_changed_CCREs)), rep('-', nrow(A549_GR_no_changed_CCREs)))


U2OS_GR_U2OS_induced <- toUpSet[toUpSet$GR_U2OS == 1 & toUpSet$ind_U2OS == 1, 1:3]
U2OS_GR_U2OS_induced <- center.bed(U2OS_GR_U2OS_induced, 500,499)
U2OS_GR_U2OS_induced_bed <- rbind(U2OS_GR_U2OS_induced, U2OS_GR_U2OS_induced)
U2OS_GR_U2OS_induced_bed$V4 <- '.'
U2OS_GR_U2OS_induced_bed$V5 <- '.'
U2OS_GR_U2OS_induced_bed$V6 <-
  c(rep('+', nrow(U2OS_GR_U2OS_induced)), rep('-', nrow(U2OS_GR_U2OS_induced)))

U2OS_GR_no_changed_CCREs <- toUpSet[toUpSet$GR_U2OS == 1 & toUpSet$ind_U2OS == 0, 1:3]
U2OS_GR_no_changed_CCREs <- center.bed(U2OS_GR_no_changed_CCREs, 500,499)
U2OS_GR_no_changed_bed <- rbind(U2OS_GR_no_changed_CCREs, U2OS_GR_no_changed_CCREs)
U2OS_GR_no_changed_bed$V4 <- '.'
U2OS_GR_no_changed_bed$V5 <- '.'
U2OS_GR_no_changed_bed$V6 <-
```

```
  c(rep('+', nrow(U2OS_GR_no_changed_CCREs)), rep('-', nrow(U2OS_GR_no_changed_CCREs)))

# Loading bigwigs
A0F <- load.bigWig(paste0(ssBigwigPath, 'A549_WT_0dex_merged_subsampled_fwd.bw'))
A0R <- load.bigWig(paste0(ssBigwigPath, 'A549_WT_0dex_merged_subsampled_rev.bw'))

A100F <- load.bigWig(paste0(ssBigwigPath, 'A549_WT_100dex_merged_subsampled_fwd.bw'))
A100R <- load.bigWig(paste0(ssBigwigPath, 'A549_WT_100dex_merged_subsampled_rev.bw'))

U0F <- load.bigWig(paste0(ssBigwigPath, 'U2OS_WT_0dex_merged_subsampled_fwd.bw'))
U0R <- load.bigWig(paste0(ssBigwigPath, 'U2OS_WT_0dex_merged_subsampled_rev.bw'))

U100F <- load.bigWig(paste0(ssBigwigPath, 'U2OS_WT_100dex_merged_subsampled_fwd.bw'))
U100R <- load.bigWig(paste0(ssBigwigPath, 'U2OS_WT_100dex_merged_subsampled_rev.bw'))

# Making metaprofiles

A549_GR_A549_induced_0_meta_F <-
  metaprofile.bigWig(A549_GR_A549_induced_bed, A0F, A0R, 10)
A549_GR_A549_induced_0_meta_R <-
  metaprofile.bigWig(A549_GR_A549_induced_bed, A0R, A0F, 10)
A549_GR_A549_induced_100_meta_F <-
  metaprofile.bigWig(A549_GR_A549_induced_bed, A100F, A100R, 10)
A549_GR_A549_induced_100_meta_R <-
  metaprofile.bigWig(A549_GR_A549_induced_bed, A100R, A100F, 10)

A549_GR_no_changed_0_meta_F <-
  metaprofile.bigWig(A549_GR_no_changed_bed, A0F, A0R, 10)
A549_GR_no_changed_0_meta_R <-
  metaprofile.bigWig(A549_GR_no_changed_bed, A0R, A0F, 10)
A549_GR_no_changed_100_meta_F <-
  metaprofile.bigWig(A549_GR_no_changed_bed, A100F, A100R, 10)
A549_GR_no_changed_100_meta_R <-
  metaprofile.bigWig(A549_GR_no_changed_bed, A100R, A100F, 10)

U2OS_GR_U2OS_induced_0_meta_F <-
  metaprofile.bigWig(U2OS_GR_U2OS_induced_bed, U0F, U0R, 10)
U2OS_GR_U2OS_induced_0_meta_R <-
  metaprofile.bigWig(U2OS_GR_U2OS_induced_bed, U0R, U0F, 10)
U2OS_GR_U2OS_induced_100_meta_F <-
  metaprofile.bigWig(U2OS_GR_U2OS_induced_bed, U100F, U100R, 10)
U2OS_GR_U2OS_induced_100_meta_R <-
  metaprofile.bigWig(U2OS_GR_U2OS_induced_bed, U100R, U100F, 10)

U2OS_GR_no_changed_0_meta_F <-
  metaprofile.bigWig(U2OS_GR_no_changed_bed, U0F, U0R, 10)
U2OS_GR_no_changed_0_meta_R <-
  metaprofile.bigWig(U2OS_GR_no_changed_bed, U0R, U0F, 10)
U2OS_GR_no_changed_100_meta_F <-
  metaprofile.bigWig(U2OS_GR_no_changed_bed, U100F, U100R, 10)
U2OS_GR_no_changed_100_meta_R <-
  metaprofile.bigWig(U2OS_GR_no_changed_bed, U100R, U100F, 10)
```

```r
#Plotting

x = 1:100*10


par(mfrow=c(2,2))
plot.metaprofile(
  A549_GR_no_changed_0_meta_F,
  minus.profile=A549_GR_no_changed_0_meta_R,
  X0=1, ylim=c(-1,1), draw.error = F, col = c('#515151', '#515151')
  )
lines(x, A549_GR_A549_induced_0_meta_F$middle, col="#7570b3")
lines(x, -1*A549_GR_A549_induced_0_meta_R$middle, col="#7570b3")

plot.metaprofile(
  A549_GR_no_changed_100_meta_F,
  minus.profile=A549_GR_no_changed_100_meta_R,
  X0=1, ylim=c(-1,1), draw.error = F, col = c('#515151', '#515151'))
lines(x, A549_GR_A549_induced_100_meta_F$middle, col="#7570b3")
lines(x, -1*A549_GR_A549_induced_100_meta_R$middle, col="#7570b3")

plot.metaprofile(
  U2OS_GR_no_changed_0_meta_F,
  minus.profile=U2OS_GR_no_changed_0_meta_R,
  X0=1, ylim=c(-1,1), draw.error = F, col = c('#515151', '#515151'))
lines(x, U2OS_GR_U2OS_induced_0_meta_F$middle, col="#1b9e77")
lines(x, -1*U2OS_GR_U2OS_induced_0_meta_R$middle, col="#1b9e77")


plot.metaprofile(
  U2OS_GR_no_changed_100_meta_F,
  minus.profile=U2OS_GR_no_changed_0_meta_R,
  X0=1,ylim=c(-1,1), draw.error = F, col = c('#515151', '#515151'))
lines(x, U2OS_GR_U2OS_induced_100_meta_F$middle, col="#1b9e77")
lines(x, -1*U2OS_GR_U2OS_induced_100_meta_R$middle, col="#1b9e77")
```



Finding closest genes to GORs with and without changed CCRE transcription

```
# Names of differentially expressed genes
genes_A549_diff <- row.names(
  geneData[geneData$A549_wt_100vs0dex_fdr < 0.05 & geneData$A549_wt_100vs0dex_log2 > 0, ])
genes_U2OS_diff <- row.names(
  geneData[geneData$U2OS_wt_100vs0dex_fdr < 0.05 & geneData$U2OS_wt_100vs0dex_log2 > 0, ])

#Promoter coordinates for differnetially expressed genes
A549_diff_promoters <- promA549[
  promA549$V4 %in% genes_A549_diff,]
U2OS_diff_promoters <- promU2OS[
  promU2OS$V4 %in% genes_U2OS_diff,]

#Sorting bed files for GORs
A549_GR_A549_induced <- bedtoolsr::bt.sort(A549_GR_A549_induced)
A549_GR_no_changed_CCREs <- bedtoolsr::bt.sort(A549_GR_no_changed_CCREs)
U2OS_GR_U2OS_induced <- bedtoolsr::bt.sort(U2OS_GR_U2OS_induced)
U2OS_GR_no_changed_CCREs <- bedtoolsr::bt.sort(U2OS_GR_no_changed_CCREs)

#Finding promoters closest to GORs
A549_CCRE_closest <-
  bedtoolsr::bt.closest(d = T, a = A549_GR_A549_induced, b = A549_diff_promoters)
A549_GR_closest <-
```

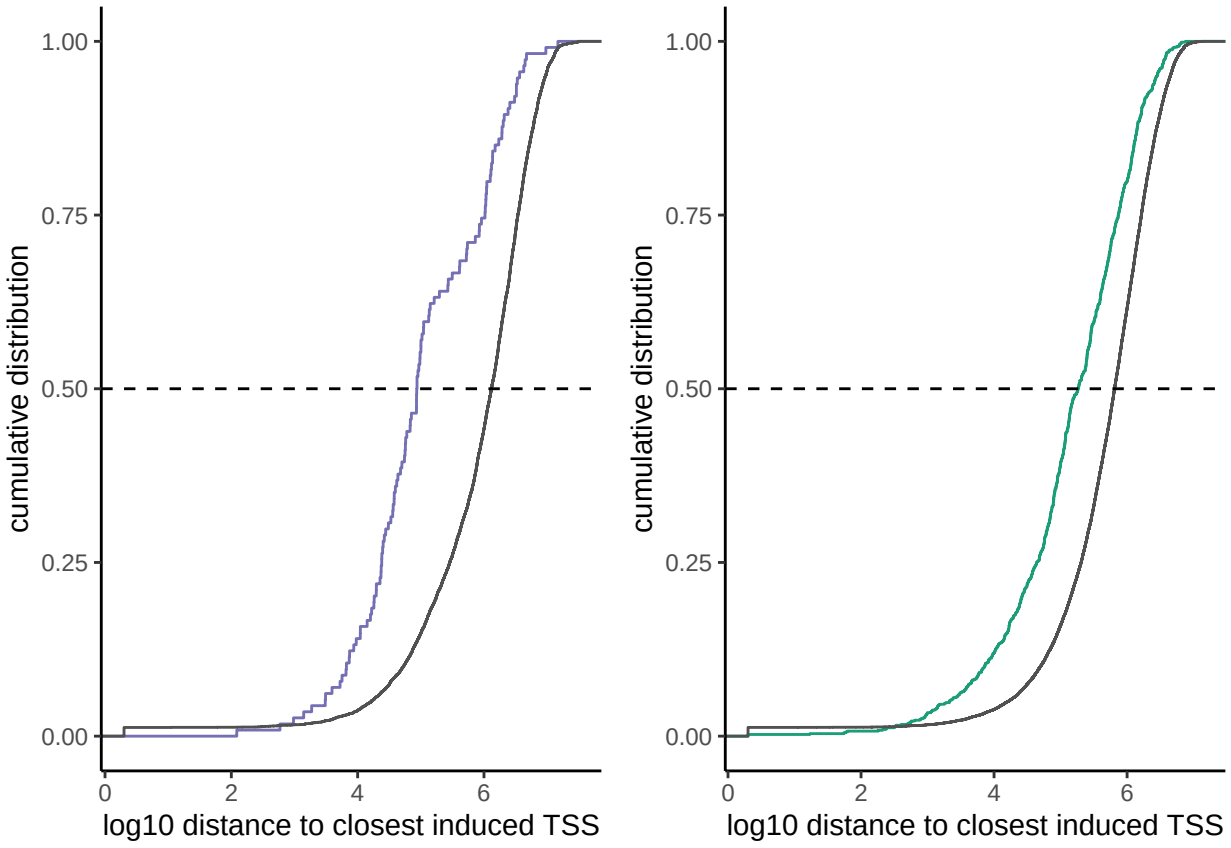
```
  bedtoolsr::bt.closest(d = T, a = A549_GR_no_changed_CCREs, b = A549_diff_promoters)
U2OS_CCRE_closest <-
  bedtoolsr::bt.closest(d = T, a = U2OS_GR_U2OS_induced, b = U2OS_diff_promoters)
U2OS_GR_closest <-
  bedtoolsr::bt.closest(d = T, a = U2OS_GR_no_changed_CCREs, b = U2OS_diff_promoters)

#Plotting CDFs
p1  <-  ggplot() +
  stat_ecdf(data = A549_CCRE_closest, aes(x=log10(2+V10)), color='#7570b3') +
  stat_ecdf(data = A549_GR_closest, aes(x=log10(2+V10)), color='#515151') +
  gg_options + xlab('log10 distance to closest induced TSS') +
  ylab('cumulative distribution') + geom_hline(yintercept = 0.5, linetype='dashed')

 wilcox.test(A549_GR_closest$V10, A549_CCRE_closest$V10)
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  A549_GR_closest$V10 and A549_CCRE_closest$V10
## W = 615193, p-value < 2.2e-16
## alternative hypothesis: true location shift is not equal to 0
```

```
p2  <-  ggplot() +
  stat_ecdf(data = U2OS_CCRE_closest, aes(x=log10(2+V10)), color='#1b9e77') +
  stat_ecdf(data = U2OS_GR_closest, aes(x=log10(2+V10)), color='#515151') +
  gg_options + xlab('log10 distance to closest induced TSS') +
  ylab('cumulative distribution') + geom_hline(yintercept = 0.5, linetype='dashed')

 wilcox.test(U2OS_GR_closest$V10, U2OS_CCRE_closest$V10)
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  U2OS_GR_closest$V10 and U2OS_CCRE_closest$V10
## W = 13425855, p-value < 2.2e-16
## alternative hypothesis: true location shift is not equal to 0
```

```
 p3 <- grid.arrange(p1, p2, nrow = 1)
```



Running HOMER to find motif enrichment in transcribed vs not GR sites

```
bedtoolsr::bt.getfasta(
  fi = '../Manuscript_data/hg38.fa',
  bed = A549_GR_A549_induced,
  fo = 'A549_GR_A549_induced.fa'
  )
```

```
## data frame with 0 columns and 0 rows
```

```
bedtoolsr::bt.getfasta(
  fi = '../Manuscript_data/hg38.fa',
  bed = A549_GR_no_changed_CCREs,
  fo = 'A549_GR_no_changed_CCREs.fa'
  )
```

```
## data frame with 0 columns and 0 rows
```

```
bedtoolsr::bt.getfasta(
  fi = '../Manuscript_data/hg38.fa',
  bed = U2OS_GR_U2OS_induced,
  fo = 'U2OS_GR_U2OS_induced.fa'
  )
```

```
## data frame with 0 columns and 0 rows
```

```
bedtoolsr::bt.getfasta(
  fi = '../Manuscript_data/hg38.fa',
  bed = U2OS_GR_no_changed_CCREs,
  fo = 'U2OS_GR_no_changed_CCREs.fa')
```

```
## data frame with 0 columns and 0 rows
```

```
echo PATH = $PATH:/home/emw97/homer/bin/
hoco='../Manuscript_data/HOCOMOCOv11_core_HUMAN_mono_homer_format_0.001.motif'

findMotifs.pl \
A549_GR_A549_induced.fa fasta \
A549_CCRE_vs_not/ -len 8,10,12 \
-fastaBg A549_GR_no_changed_CCREs.fa \
-mcheck $hoco -bits -nogo
-mknown $hoco

findMotifs.pl \
U2OS_GR_U2OS_induced.fa fasta \
U2OS_CCRE_vs_not/ -len 8,10,12 \
-fastaBg U2OS_GR_no_changed_CCREs.fa \
-mcheck $hoco -bits -nogo \
-mknown $hoco
```

Making metaplots of conservation and H3K27ac in transcribed vs not GR sites using Deeptools

```
computeMatrix reference-point \
-R A549_GR_A549_induced.bed A549_GR_no_changed_CCREs.bed \
-S GR_ChIPseq_A549_0nMdex.bw GR_ChIPseq_A549_100nMdex.bw \
A549_ATAC_EtOH.pval.bw H3K27ac_0hr.pval.bw \
H3K27ac_1hr.pval.bw H3K27ac_12hr.pval.bw  hg38.phastCons30way.bw \
-out A549_H3K27ac_phastCons_GR_txn_not_1kb.computeMatrix.gz \
--referencePoint center --missingDataAsZero \
--upstream 500 --downstream 500  -p 8

computeMatrix reference-point \
-R U2OS_GR_U2OS_induced.bed U2OS_GR_no_changed_CCREs.bed \
-S GR_ChIPseq_U2OS-hGR_0nMdex.bw GR_ChIPseq_U2OS-hGR_100nMdex.bw \
U2OS_ATAC_EtOH.pval.bw hg38.phastCons30way.bw \
-out U2OS_phastCons_GR_txn_not_1kb.computeMatrix.gz \
--referencePoint center -missingDataAsZero \
--upstream 500 --downstream 500 - -p 8


plotProfile \
--matrixFile A549_H3K27ac_phastCons_GR_txn_not_1kb.computeMatrix.gz \
--outFileName A549_H3K27ac_phastCons_GR_txn_not_1kb.metaplot.pdf \
--refPointLabel 'center' --yMax 150 150 75 300 300 300 0.5 --yMin 0
```

```
plotProfile \
--matrixFile U2OS_phastCons_GR_txn_not_1kb.computeMatrix.gz \
--outFileName U2OS_phastCons_GR_txn_not_1kb.metaplot.pdf \
--refPointLabel 'center' --yMax 150 150 75 0.5 --yMin 0
```

For CCREs that are induced by do not bind GR, seeing what fraction of them are located within induced genes (and so their induction may be an artifact)

```
ind_A549_no_GR <- toUpSet[toUpSet$GR_A549 == 0 & toUpSet$ind_A549 == 1, 1:3]
ind_U2OS_no_GR <- toUpSet[toUpSet$GR_U2OS == 0 & toUpSet$ind_U2OS == 1, 1:3]

ind_A549_no_GR_genes <- unique(
  bedtoolsr::bt.intersect(
    wb = T, a = ind_A549_no_GR,
    b = '../Manuscript_data/hg38_refseq.bed'
    )
  [,c(1,2,3,7)]
  )
ind_U2OS_no_GR_genes <- unique(
  bedtoolsr::bt.intersect(
    wb = T, a = ind_U2OS_no_GR,
    b = '../Manuscript_data/hg38_refseq.bed'
    )
  [,c(1,2,3,7)]
  )

genes_A549_diff <-
  row.names(
    geneData[geneData$A549_wt_100vs0dex_log2 > 0 &
               geneData$A549_wt_100vs0dex_fdr < 0.05,]
    )
genes_U2OS_diff <-
  row.names(
    geneData[geneData$U2OS_wt_100vs0dex_log2 > 0 &
               geneData$U2OS_wt_100vs0dex_fdr < 0.05,]
    )

A549_ind_in_genes <- ind_A549_no_GR_genes[ind_A549_no_GR_genes$V7 %in% genes_A549_diff,]
U2OS_ind_in_genes <- ind_U2OS_no_GR_genes[ind_U2OS_no_GR_genes$V7 %in% genes_U2OS_diff,]

print(c('Number of A549 ind CCREs w/o GR', nrow(ind_A549_no_GR)))
```

```
## [1] "Number of A549 ind CCREs w/o GR" "79"
```

```
print(c('Number of A549 ind CCREs w/o GR within induced genes', nrow(A549_ind_in_genes)))
```

```
## [1] "Number of A549 ind CCREs w/o GR within induced genes"
## [2] "32"
```

```r
print(c('Number of U2OS ind CCREs w/o GR', nrow(ind_U2OS_no_GR)))
```

```
## [1] "Number of U2OS ind CCREs w/o GR" "349"
```

```r
print(c('Number of U2OS ind CCREs w/o GR within induced genes', nrow(U2OS_ind_in_genes)))
```

```
## [1] "Number of U2OS ind CCREs w/o GR within induced genes"
## [2] "115"
```