capture_coverage_analysis

Jacob Green

2023-03-14

R Markdown

```
library(ggplot2)
library(grid)
library(plyr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:plyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarize

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(scales)
library(zoo)
```

```
##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':
##
##     combine
```

```
library(grid)
```

```
setwd("/home/jgreen/EAGER_OBJ1a/04_coverage_analysis/01_genome_region/")
# initialize an empty list to store ggplot objects
plot_list <- list()
```

```r
files <- c("Capture1_B3.hist", "Capture2_B3.hist", "Capture3_B3.hist", "Capture4_B3.hist")

for (file in files) {
  print(file)
  files_list <- list.files(pattern=file)
  print(files_list)

  files <- c(paste0(file, ".AllCDS.all.split.txt"), paste0(file, ".AllExon.all.split.txt"), paste0(file
  print(files)

  labs <- c("CDS","Exon","Gene","UTR")

  cov <- list()
  for (i in 1:length(files)) {
    cov[[i]] <- read.table(files[i])[,c(2,5)]
    cov_cumul=1-cumsum(cov[[i]][,2])
    cov[[i]]$cov_cumul <- c(1,cov_cumul[-length(cov_cumul)])
    cov[[i]]$sample=labs[i]
  }

  cov_df=do.call("rbind",cov)
  names(cov_df)[1:2]=c("depth","fraction")

  pcbPalette <- c("#009E73" ,"#D55E00","#CC79A7","#56B4E9")

  p1 <- ggplot(cov_df, aes(x= depth, y=cov_cumul, color=sample))  + xlim(0,25)+
    scale_alpha(guide = 'none') +
    geom_line(size=1.5)+
    #geom_segment(aes(x=20, y=0, xend=20, yend=1, color="red"))+
    scale_color_manual(values=pcbPalette) +
    scale_fill_manual(values=pcbPalette) +
    ggtitle(file)+
    ylab("% of Bases > Depth")+
    xlab("Depth")+
    theme_bw() +
    theme(plot.title = element_text(size = 12, face = "bold",hjust = 0.5)) +
    theme(legend.title = element_blank()) +
    theme(legend.position=c(0.50,0.75))

  plot_list[[length(plot_list)+1]] <- p1

  png(filename=paste0("Figure2_", file, ".png"), type="cairo",units="px", width=5600,
      height=3000, res=600, bg="transparent")

  print(p1)

  dev.off()

}
```

```
## [1] "Capture1_B3.hist"
## [1] "Capture1_B3.hist.AllCDS.all.split.txt"
## [2] "Capture1_B3.hist.AllExon.all.split.txt"
```

```
## [3] "Capture1_B3.hist.AllGene.all.split.txt"
## [4] "Capture1_B3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture1_B3.hist.png"
## [1] "Capture1_B3.hist.AllCDS.all.split.txt"
## [2] "Capture1_B3.hist.AllExon.all.split.txt"
## [3] "Capture1_B3.hist.AllGene.all.split.txt"
## [4] "Capture1_B3.hist.AllUTR.all.split.txt"

## Warning: Removed 4937 row(s) containing missing values (geom_path).

## [1] "Capture2_B3.hist"
## [1] "Capture2_B3.hist.AllCDS.all.split.txt"
## [2] "Capture2_B3.hist.AllExon.all.split.txt"
## [3] "Capture2_B3.hist.AllGene.all.split.txt"
## [4] "Capture2_B3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture2_B3.hist.png"
## [1] "Capture2_B3.hist.AllCDS.all.split.txt"
## [2] "Capture2_B3.hist.AllExon.all.split.txt"
## [3] "Capture2_B3.hist.AllGene.all.split.txt"
## [4] "Capture2_B3.hist.AllUTR.all.split.txt"

## Warning: Removed 5091 row(s) containing missing values (geom_path).

## [1] "Capture3_B3.hist"
## [1] "Capture3_B3.hist.AllCDS.all.split.txt"
## [2] "Capture3_B3.hist.AllExon.all.split.txt"
## [3] "Capture3_B3.hist.AllGene.all.split.txt"
## [4] "Capture3_B3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture3_B3.hist.png"
## [1] "Capture3_B3.hist.AllCDS.all.split.txt"
## [2] "Capture3_B3.hist.AllExon.all.split.txt"
## [3] "Capture3_B3.hist.AllGene.all.split.txt"
## [4] "Capture3_B3.hist.AllUTR.all.split.txt"

## Warning: Removed 2209 row(s) containing missing values (geom_path).

## [1] "Capture4_B3.hist"
## [1] "Capture4_B3.hist.AllCDS.all.split.txt"
## [2] "Capture4_B3.hist.AllExon.all.split.txt"
## [3] "Capture4_B3.hist.AllGene.all.split.txt"
## [4] "Capture4_B3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture4_B3.hist.png"
## [1] "Capture4_B3.hist.AllCDS.all.split.txt"
## [2] "Capture4_B3.hist.AllExon.all.split.txt"
## [3] "Capture4_B3.hist.AllGene.all.split.txt"
## [4] "Capture4_B3.hist.AllUTR.all.split.txt"

## Warning: Removed 3176 row(s) containing missing values (geom_path).
# combine all the ggplot objects in the list using grid.arrange()
pcombined <- grid.arrange(grobs = plot_list, ncol = 2, top=textGrob("Capture B3 Read Depth across genome

## Warning: Removed 4937 row(s) containing missing values (geom_path).

## Warning: Removed 5091 row(s) containing missing values (geom_path).

## Warning: Removed 2209 row(s) containing missing values (geom_path).

## Warning: Removed 3176 row(s) containing missing values (geom_path).
```
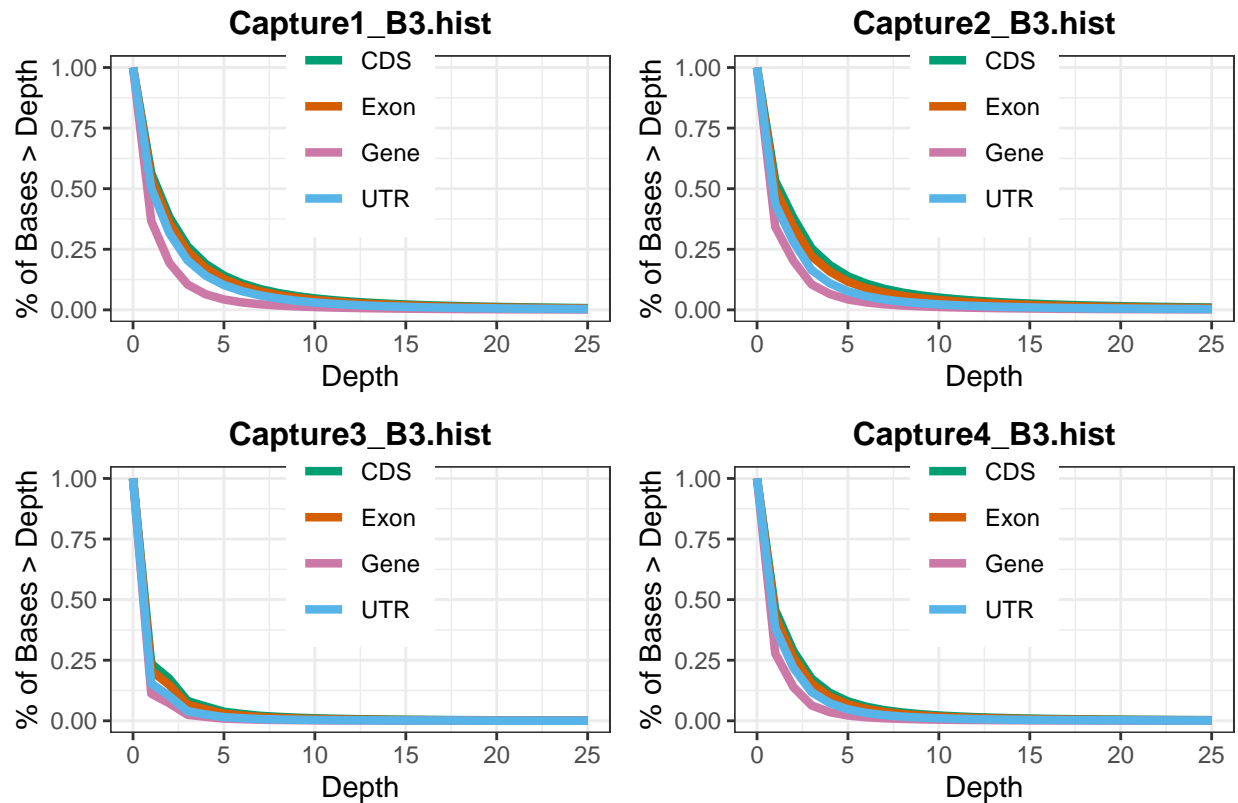
Capture B3 Read Depth across genome


```
#save the combined plot as PNG file
ggsave(file="Figure2_allcapture_B3.png", pcombined)
```

```
## Saving 6.5 x 4.5 in image
```

```
setwd("/home/jgreen/EAGER_OBJ1a/04_coverage_analysis/01_genome_region/")
# initialize an empty list to store ggplot objects
plot_list <- list()

files <- c("Capture1_B4.hist", "Capture2_B4.hist", "Capture3_B4.hist", "Capture4_B4.hist")

for (file in files) {
  print(file)
  files_list <- list.files(pattern=file)
  print(files_list)

  files <- c(paste0(file, ".AllCDS.all.split.txt"), paste0(file, ".AllExon.all.split.txt"), paste0(file
  print(files)

  labs <- c("CDS","Exon","Gene","UTR")

  cov <- list()
  for (i in 1:length(files)) {
    cov[[i]] <- read.table(files[i])[,c(2,5)]
    cov_cumul=1-cumsum(cov[[i]][,2])
    cov[[i]]$cov_cumul <- c(1,cov_cumul[-length(cov_cumul)])
    cov[[i]]$sample=labs[i]
```

```r
  }

  cov_df=do.call("rbind",cov)
  names(cov_df)[1:2]=c("depth","fraction")

  pcbPalette <- c("#009E73" ,"#D55E00","#CC79A7","#56B4E9")

  p1 <- ggplot(cov_df, aes(x= depth, y=cov_cumul, color=sample))  + xlim(0,25)+
    scale_alpha(guide = 'none') +
    geom_line(size=1.5)+
    #geom_segment(aes(x=20, y=0, xend=20, yend=1, color="red"))+
    scale_color_manual(values=pcbPalette) +
    scale_fill_manual(values=pcbPalette) +
    ggtitle(file)+
    ylab("% of Bases > Depth")+
    xlab("Depth")+
    theme_bw() +
    theme(plot.title = element_text(size = 12, face = "bold",hjust = 0.5)) +
    theme(legend.title = element_blank()) +
    theme(legend.position=c(0.50,0.75))

  plot_list[[length(plot_list)+1]] <- p1

  png(filename=paste0("Figure2_", file, ".png"), type="cairo",units="px", width=5600,
      height=3000, res=600, bg="transparent")

  print(p1)

  dev.off()

}
```

```
## [1] "Capture1_B4.hist"
## [1] "Capture1_B4.hist.AllCDS.all.split.txt"
## [2] "Capture1_B4.hist.AllExon.all.split.txt"
## [3] "Capture1_B4.hist.AllGene.all.split.txt"
## [4] "Capture1_B4.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture1_B4.hist.png"
## [1] "Capture1_B4.hist.AllCDS.all.split.txt"
## [2] "Capture1_B4.hist.AllExon.all.split.txt"
## [3] "Capture1_B4.hist.AllGene.all.split.txt"
## [4] "Capture1_B4.hist.AllUTR.all.split.txt"

## Warning: Removed 3789 row(s) containing missing values (geom_path).

## [1] "Capture2_B4.hist"
## [1] "Capture2_B4.hist.AllCDS.all.split.txt"
## [2] "Capture2_B4.hist.AllExon.all.split.txt"
## [3] "Capture2_B4.hist.AllGene.all.split.txt"
## [4] "Capture2_B4.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture2_B4.hist.png"
## [1] "Capture2_B4.hist.AllCDS.all.split.txt"
## [2] "Capture2_B4.hist.AllExon.all.split.txt"
## [3] "Capture2_B4.hist.AllGene.all.split.txt"
## [4] "Capture2_B4.hist.AllUTR.all.split.txt"
```

```
## Warning: Removed 4642 row(s) containing missing values (geom_path).

## [1] "Capture3_B4.hist"
## [1] "Capture3_B4.hist.AllCDS.all.split.txt"
## [2] "Capture3_B4.hist.AllExon.all.split.txt"
## [3] "Capture3_B4.hist.AllGene.all.split.txt"
## [4] "Capture3_B4.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture3_B4.hist.png"
## [1] "Capture3_B4.hist.AllCDS.all.split.txt"
## [2] "Capture3_B4.hist.AllExon.all.split.txt"
## [3] "Capture3_B4.hist.AllGene.all.split.txt"
## [4] "Capture3_B4.hist.AllUTR.all.split.txt"

## Warning: Removed 4938 row(s) containing missing values (geom_path).

## [1] "Capture4_B4.hist"
## [1] "Capture4_B4.hist.AllCDS.all.split.txt"
## [2] "Capture4_B4.hist.AllExon.all.split.txt"
## [3] "Capture4_B4.hist.AllGene.all.split.txt"
## [4] "Capture4_B4.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture4_B4.hist.png"
## [1] "Capture4_B4.hist.AllCDS.all.split.txt"
## [2] "Capture4_B4.hist.AllExon.all.split.txt"
## [3] "Capture4_B4.hist.AllGene.all.split.txt"
## [4] "Capture4_B4.hist.AllUTR.all.split.txt"

## Warning: Removed 5141 row(s) containing missing values (geom_path).
```
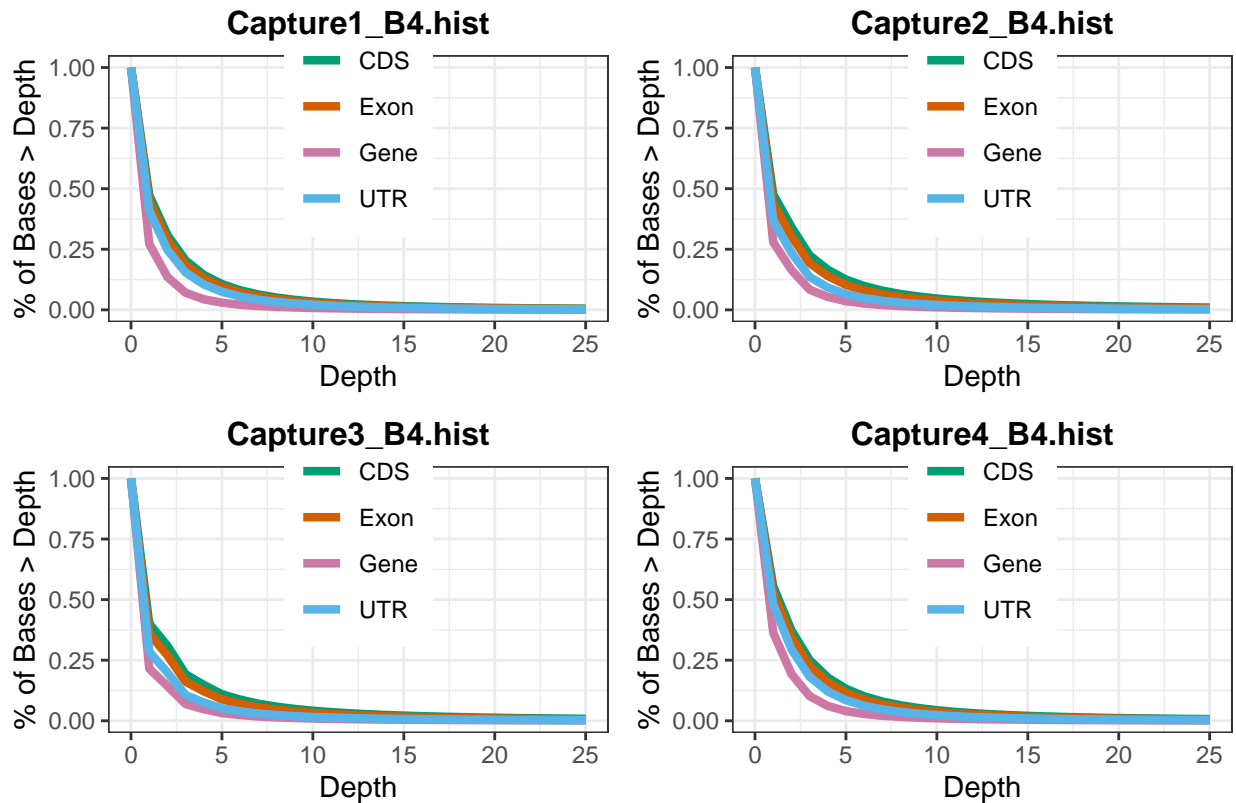
```r
# combine all the ggplot objects in the list using grid.arrange()
pcombined <- grid.arrange(grobs = plot_list, ncol = 2, top=textGrob("Capture B4 Read Depth across genom
```

```
## Warning: Removed 3789 row(s) containing missing values (geom_path).

## Warning: Removed 4642 row(s) containing missing values (geom_path).

## Warning: Removed 4938 row(s) containing missing values (geom_path).

## Warning: Removed 5141 row(s) containing missing values (geom_path).
```

Capture B4 Read Depth across genome


```
#save the combined plot as PNG file
ggsave(file="Figure2_allcapture_B4.png", pcombined)
```

```
## Saving 6.5 x 4.5 in image
```

```
setwd("/home/jgreen/EAGER_OBJ1a/04_coverage_analysis/01_genome_region/")
# initialize an empty list to store ggplot objects
plot_list <- list()

files <- c("Capture1_G3.hist", "Capture2_G3.hist", "Capture3_G3.hist", "Capture4_G3.hist")

for (file in files) {
  print(file)
  files_list <- list.files(pattern=file)
  print(files_list)

  files <- c(paste0(file, ".AllCDS.all.split.txt"), paste0(file, ".AllExon.all.split.txt"), paste0(file
  print(files)

  labs <- c("CDS","Exon","Gene","UTR")

  cov <- list()
  for (i in 1:length(files)) {
    cov[[i]] <- read.table(files[i])[,c(2,5)]
    cov_cumul=1-cumsum(cov[[i]][,2])
    cov[[i]]$cov_cumul <- c(1,cov_cumul[-length(cov_cumul)])
    cov[[i]]$sample=labs[i]
```

```
  }

  cov_df=do.call("rbind",cov)
  names(cov_df)[1:2]=c("depth","fraction")

  pcbPalette <- c("#009E73" ,"#D55E00","#CC79A7","#56B4E9")

  p1 <- ggplot(cov_df, aes(x= depth, y=cov_cumul, color=sample))  + xlim(0,25)+
    scale_alpha(guide = 'none') +
    geom_line(size=1.5)+
    #geom_segment(aes(x=20, y=0, xend=20, yend=1, color="red"))+
    scale_color_manual(values=pcbPalette) +
    scale_fill_manual(values=pcbPalette) +
    ggtitle(file)+
    ylab("% of Bases > Depth")+
    xlab("Depth")+
    theme_bw() +
    theme(plot.title = element_text(size = 12, face = "bold",hjust = 0.5)) +
    theme(legend.title = element_blank()) +
    theme(legend.position=c(0.50,0.75))

  plot_list[[length(plot_list)+1]] <- p1

  png(filename=paste0("Figure2_", file, ".png"), type="cairo",units="px", width=5600,
      height=3000, res=600, bg="transparent")

  print(p1)

  dev.off()

}
```

```
## [1] "Capture1_G3.hist"
## [1] "Capture1_G3.hist.AllCDS.all.split.txt"
## [2] "Capture1_G3.hist.AllExon.all.split.txt"
## [3] "Capture1_G3.hist.AllGene.all.split.txt"
## [4] "Capture1_G3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture1_G3.hist.png"
## [1] "Capture1_G3.hist.AllCDS.all.split.txt"
## [2] "Capture1_G3.hist.AllExon.all.split.txt"
## [3] "Capture1_G3.hist.AllGene.all.split.txt"
## [4] "Capture1_G3.hist.AllUTR.all.split.txt"

## Warning: Removed 6111 row(s) containing missing values (geom_path).

## [1] "Capture2_G3.hist"
## [1] "Capture2_G3.hist.AllCDS.all.split.txt"
## [2] "Capture2_G3.hist.AllExon.all.split.txt"
## [3] "Capture2_G3.hist.AllGene.all.split.txt"
## [4] "Capture2_G3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture2_G3.hist.png"
## [1] "Capture2_G3.hist.AllCDS.all.split.txt"
## [2] "Capture2_G3.hist.AllExon.all.split.txt"
## [3] "Capture2_G3.hist.AllGene.all.split.txt"
## [4] "Capture2_G3.hist.AllUTR.all.split.txt"
```

```
## Warning: Removed 8517 row(s) containing missing values (geom_path).

## [1] "Capture3_G3.hist"
## [1] "Capture3_G3.hist.AllCDS.all.split.txt"
## [2] "Capture3_G3.hist.AllExon.all.split.txt"
## [3] "Capture3_G3.hist.AllGene.all.split.txt"
## [4] "Capture3_G3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture3_G3.hist.png"
## [1] "Capture3_G3.hist.AllCDS.all.split.txt"
## [2] "Capture3_G3.hist.AllExon.all.split.txt"
## [3] "Capture3_G3.hist.AllGene.all.split.txt"
## [4] "Capture3_G3.hist.AllUTR.all.split.txt"

## Warning: Removed 5838 row(s) containing missing values (geom_path).

## [1] "Capture4_G3.hist"
## [1] "Capture4_G3.hist.AllCDS.all.split.txt"
## [2] "Capture4_G3.hist.AllExon.all.split.txt"
## [3] "Capture4_G3.hist.AllGene.all.split.txt"
## [4] "Capture4_G3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture4_G3.hist.png"
## [1] "Capture4_G3.hist.AllCDS.all.split.txt"
## [2] "Capture4_G3.hist.AllExon.all.split.txt"
## [3] "Capture4_G3.hist.AllGene.all.split.txt"
## [4] "Capture4_G3.hist.AllUTR.all.split.txt"

## Warning: Removed 3048 row(s) containing missing values (geom_path).
```
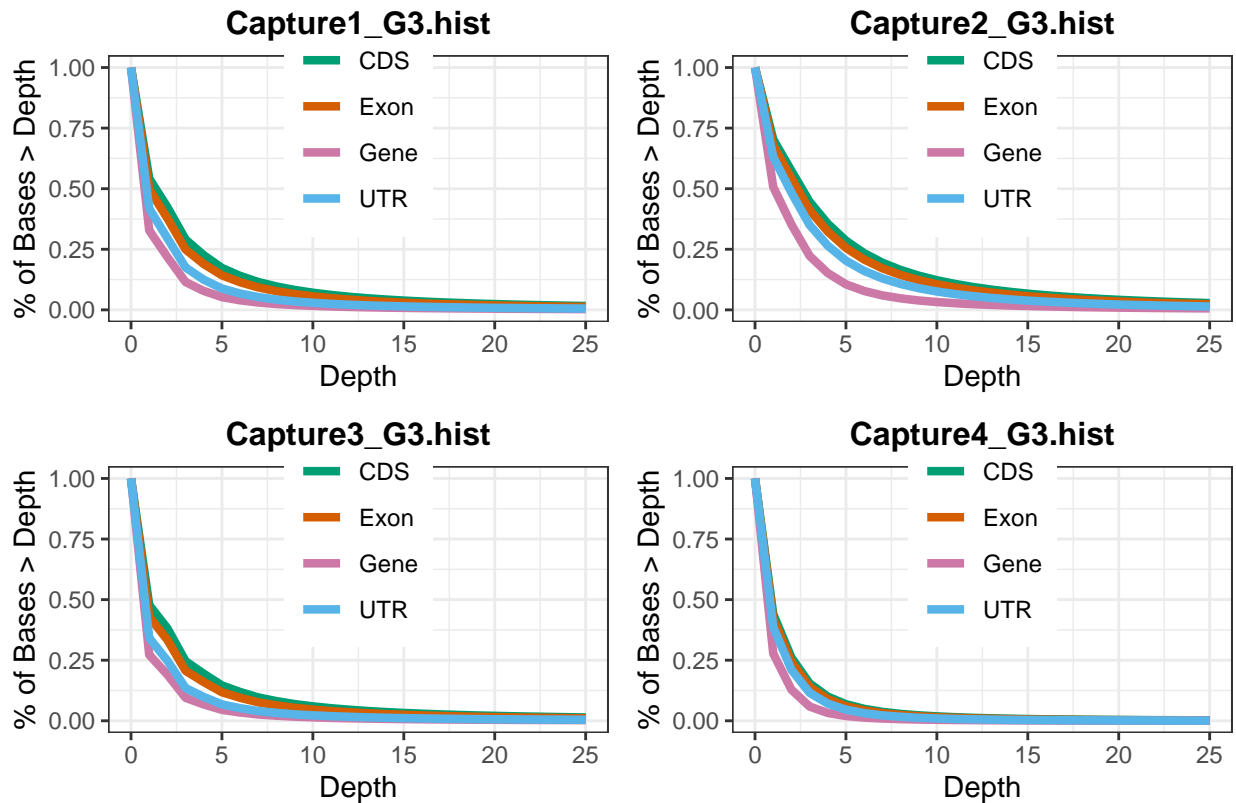
```r
# combine all the ggplot objects in the list using grid.arrange()
pcombined <- grid.arrange(grobs = plot_list, ncol = 2, top=textGrob("Capture B4 Read Depth across genom
```

```
## Warning: Removed 6111 row(s) containing missing values (geom_path).

## Warning: Removed 8517 row(s) containing missing values (geom_path).

## Warning: Removed 5838 row(s) containing missing values (geom_path).

## Warning: Removed 3048 row(s) containing missing values (geom_path).
```

Capture B4 Read Depth across genome


```r
#save the combined plot as PNG file
ggsave(file="Figure2_allcapture_G3.png", pcombined)
```

```
## Saving 6.5 x 4.5 in image
```

```r
setwd("/home/jgreen/EAGER_OBJ1a/04_coverage_analysis/01_genome_region/")
# initialize an empty list to store ggplot objects
plot_list <- list()

files <- c("Capture1_G5.hist", "Capture2_G5.hist", "Capture3_G5.hist", "Capture4_G5.hist")

for (file in files) {
  print(file)
  files_list <- list.files(pattern=file)
  print(files_list)

  files <- c(paste0(file, ".AllCDS.all.split.txt"), paste0(file, ".AllExon.all.split.txt"), paste0(file
  print(files)

  labs <- c("CDS","Exon","Gene","UTR")

  cov <- list()
  for (i in 1:length(files)) {
    cov[[i]] <- read.table(files[i])[,c(2,5)]
    cov_cumul=1-cumsum(cov[[i]][,2])
    cov[[i]]$cov_cumul <- c(1,cov_cumul[-length(cov_cumul)])
    cov[[i]]$sample=labs[i]
```

```r
  }

  cov_df=do.call("rbind",cov)
  names(cov_df)[1:2]=c("depth","fraction")

  pcbPalette <- c("#009E73" ,"#D55E00","#CC79A7","#56B4E9")

  p1 <- ggplot(cov_df, aes(x= depth, y=cov_cumul, color=sample))  + xlim(0,25)+
    scale_alpha(guide = 'none') +
    geom_line(size=1.5)+
    #geom_segment(aes(x=20, y=0, xend=20, yend=1, color="red"))+
    scale_color_manual(values=pcbPalette) +
    scale_fill_manual(values=pcbPalette) +
    ggtitle(file)+
    ylab("% of Bases > Depth")+
    xlab("Depth")+
    theme_bw() +
    theme(plot.title = element_text(size = 12, face = "bold",hjust = 0.5)) +
    theme(legend.title = element_blank()) +
    theme(legend.position=c(0.50,0.75))

  plot_list[[length(plot_list)+1]] <- p1

  png(filename=paste0("Figure2_", file, ".png"), type="cairo",units="px", width=5600,
      height=3000, res=600, bg="transparent")

  print(p1)

  dev.off()

}
```

```
## [1] "Capture1_G5.hist"
## [1] "Capture1_G5.hist.AllCDS.all.split.txt"
## [2] "Capture1_G5.hist.AllExon.all.split.txt"
## [3] "Capture1_G5.hist.AllGene.all.split.txt"
## [4] "Capture1_G5.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture1_G5.hist.png"
## [1] "Capture1_G5.hist.AllCDS.all.split.txt"
## [2] "Capture1_G5.hist.AllExon.all.split.txt"
## [3] "Capture1_G5.hist.AllGene.all.split.txt"
## [4] "Capture1_G5.hist.AllUTR.all.split.txt"

## Warning: Removed 4575 row(s) containing missing values (geom_path).

## [1] "Capture2_G5.hist"
## [1] "Capture2_G5.hist.AllCDS.all.split.txt"
## [2] "Capture2_G5.hist.AllExon.all.split.txt"
## [3] "Capture2_G5.hist.AllGene.all.split.txt"
## [4] "Capture2_G5.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture2_G5.hist.png"
## [1] "Capture2_G5.hist.AllCDS.all.split.txt"
## [2] "Capture2_G5.hist.AllExon.all.split.txt"
## [3] "Capture2_G5.hist.AllGene.all.split.txt"
## [4] "Capture2_G5.hist.AllUTR.all.split.txt"
```

```
## Warning: Removed 4206 row(s) containing missing values (geom_path).

## [1] "Capture3_G5.hist"
## [1] "Capture3_G5.hist.AllCDS.all.split.txt"
## [2] "Capture3_G5.hist.AllExon.all.split.txt"
## [3] "Capture3_G5.hist.AllGene.all.split.txt"
## [4] "Capture3_G5.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture3_G5.hist.png"
## [1] "Capture3_G5.hist.AllCDS.all.split.txt"
## [2] "Capture3_G5.hist.AllExon.all.split.txt"
## [3] "Capture3_G5.hist.AllGene.all.split.txt"
## [4] "Capture3_G5.hist.AllUTR.all.split.txt"

## Warning: Removed 4508 row(s) containing missing values (geom_path).

## [1] "Capture4_G5.hist"
## [1] "Capture4_G5.hist.AllCDS.all.split.txt"
## [2] "Capture4_G5.hist.AllExon.all.split.txt"
## [3] "Capture4_G5.hist.AllGene.all.split.txt"
## [4] "Capture4_G5.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture4_G5.hist.png"
## [1] "Capture4_G5.hist.AllCDS.all.split.txt"
## [2] "Capture4_G5.hist.AllExon.all.split.txt"
## [3] "Capture4_G5.hist.AllGene.all.split.txt"
## [4] "Capture4_G5.hist.AllUTR.all.split.txt"

## Warning: Removed 3603 row(s) containing missing values (geom_path).
```

```r
# combine all the ggplot objects in the list using grid.arrange()
pcombined <- grid.arrange(grobs = plot_list, ncol = 2, top=textGrob("Capture G4 Read Depth across genom
```

```
## Warning: Removed 4575 row(s) containing missing values (geom_path).

## Warning: Removed 4206 row(s) containing missing values (geom_path).

## Warning: Removed 4508 row(s) containing missing values (geom_path).

## Warning: Removed 3603 row(s) containing missing values (geom_path).
```

Capture G4 Read Depth across genome
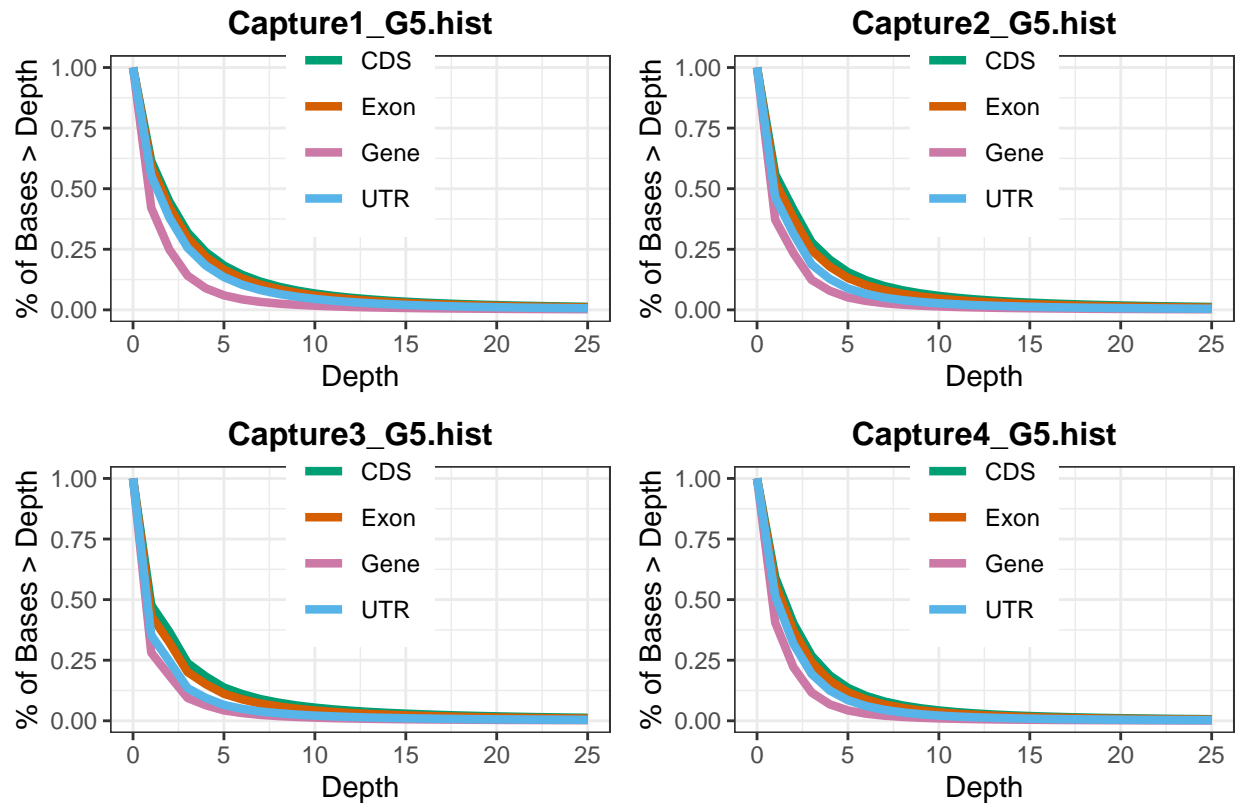

```r
#save the combined plot as PNG file
ggsave(file="Figure2_allcapture_G4.png", pcombined)
```

```
## Saving 6.5 x 4.5 in image
```

```r
setwd("/home/jgreen/EAGER_OBJ1a/04_coverage_analysis/01_genome_region/")
# initialize an empty list to store ggplot objects
plot_list <- list()

files <- c("Capture1_K3.hist", "Capture2_K3.hist", "Capture3_K3.hist", "Capture4_K3.hist")

for (file in files) {
  print(file)
  files_list <- list.files(pattern=file)
  print(files_list)

  files <- c(paste0(file, ".AllCDS.all.split.txt"), paste0(file, ".AllExon.all.split.txt"), paste0(file
  print(files)

  labs <- c("CDS","Exon","Gene","UTR")

  cov <- list()
  for (i in 1:length(files)) {
    cov[[i]] <- read.table(files[i])[,c(2,5)]
    cov_cumul=1-cumsum(cov[[i]][,2])
    cov[[i]]$cov_cumul <- c(1,cov_cumul[-length(cov_cumul)])
    cov[[i]]$sample=labs[i]
```

```
  }

  cov_df=do.call("rbind",cov)
  names(cov_df)[1:2]=c("depth","fraction")

  pcbPalette <- c("#009E73" ,"#D55E00","#CC79A7","#56B4E9")

  p1 <- ggplot(cov_df, aes(x= depth, y=cov_cumul, color=sample))  + xlim(0,25)+
    scale_alpha(guide = 'none') +
    geom_line(size=1.5)+
    #geom_segment(aes(x=20, y=0, xend=20, yend=1, color="red"))+
    scale_color_manual(values=pcbPalette) +
    scale_fill_manual(values=pcbPalette) +
    ggtitle(file)+
    ylab("% of Bases > Depth")+
    xlab("Depth")+
    theme_bw() +
    theme(plot.title = element_text(size = 12, face = "bold",hjust = 0.5)) +
    theme(legend.title = element_blank()) +
    theme(legend.position=c(0.50,0.75))

  plot_list[[length(plot_list)+1]] <- p1

  png(filename=paste0("Figure2_", file, ".png"), type="cairo",units="px", width=5600,
      height=3000, res=600, bg="transparent")

  print(p1)

  dev.off()

}
```

```
## [1] "Capture1_K3.hist"
## [1] "Capture1_K3.hist.AllCDS.all.split.txt"
## [2] "Capture1_K3.hist.AllExon.all.split.txt"
## [3] "Capture1_K3.hist.AllGene.all.split.txt"
## [4] "Capture1_K3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture1_K3.hist.png"
## [1] "Capture1_K3.hist.AllCDS.all.split.txt"
## [2] "Capture1_K3.hist.AllExon.all.split.txt"
## [3] "Capture1_K3.hist.AllGene.all.split.txt"
## [4] "Capture1_K3.hist.AllUTR.all.split.txt"

## Warning: Removed 4075 row(s) containing missing values (geom_path).

## [1] "Capture2_K3.hist"
## [1] "Capture2_K3.hist.AllCDS.all.split.txt"
## [2] "Capture2_K3.hist.AllExon.all.split.txt"
## [3] "Capture2_K3.hist.AllGene.all.split.txt"
## [4] "Capture2_K3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture2_K3.hist.png"
## [1] "Capture2_K3.hist.AllCDS.all.split.txt"
## [2] "Capture2_K3.hist.AllExon.all.split.txt"
## [3] "Capture2_K3.hist.AllGene.all.split.txt"
## [4] "Capture2_K3.hist.AllUTR.all.split.txt"
```

```
## Warning: Removed 3923 row(s) containing missing values (geom_path).

## [1] "Capture3_K3.hist"
## [1] "Capture3_K3.hist.AllCDS.all.split.txt"
## [2] "Capture3_K3.hist.AllExon.all.split.txt"
## [3] "Capture3_K3.hist.AllGene.all.split.txt"
## [4] "Capture3_K3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture3_K3.hist.png"
## [1] "Capture3_K3.hist.AllCDS.all.split.txt"
## [2] "Capture3_K3.hist.AllExon.all.split.txt"
## [3] "Capture3_K3.hist.AllGene.all.split.txt"
## [4] "Capture3_K3.hist.AllUTR.all.split.txt"

## Warning: Removed 4187 row(s) containing missing values (geom_path).

## [1] "Capture4_K3.hist"
## [1] "Capture4_K3.hist.AllCDS.all.split.txt"
## [2] "Capture4_K3.hist.AllExon.all.split.txt"
## [3] "Capture4_K3.hist.AllGene.all.split.txt"
## [4] "Capture4_K3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture4_K3.hist.png"
## [1] "Capture4_K3.hist.AllCDS.all.split.txt"
## [2] "Capture4_K3.hist.AllExon.all.split.txt"
## [3] "Capture4_K3.hist.AllGene.all.split.txt"
## [4] "Capture4_K3.hist.AllUTR.all.split.txt"

## Warning: Removed 4569 row(s) containing missing values (geom_path).
```

```r
# combine all the ggplot objects in the list using grid.arrange()
pcombined <- grid.arrange(grobs = plot_list, ncol = 2, top=textGrob("Capture K3 Read Depth across genom
```

```
## Warning: Removed 4075 row(s) containing missing values (geom_path).

## Warning: Removed 3923 row(s) containing missing values (geom_path).

## Warning: Removed 4187 row(s) containing missing values (geom_path).

## Warning: Removed 4569 row(s) containing missing values (geom_path).
```

Capture K3 Read Depth across genome
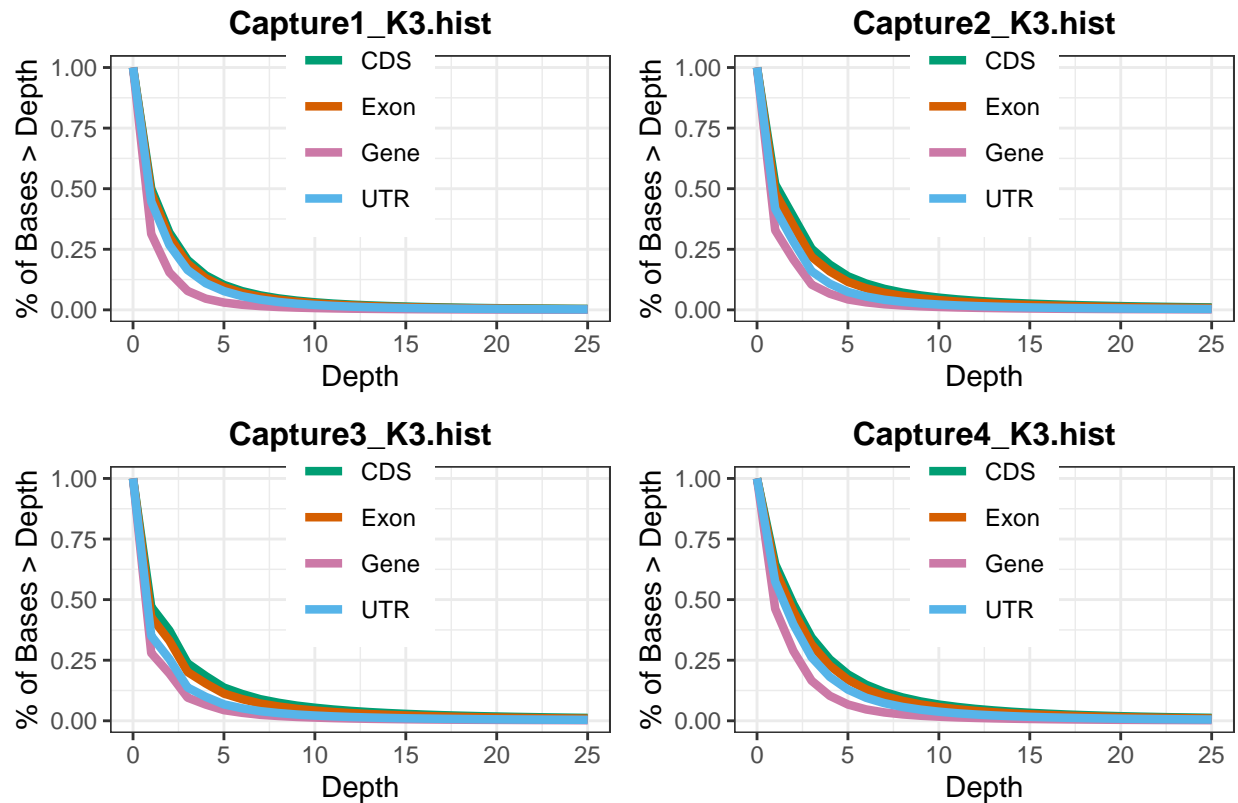
```r
#save the combined plot as PNG file
ggsave(file="Figure2_allcapture_K3.png", pcombined)
```

```
## Saving 6.5 x 4.5 in image
```

```r
setwd("/home/jgreen/EAGER_OBJ1a/04_coverage_analysis/01_genome_region/")
# initialize an empty list to store ggplot objects
plot_list <- list()

files <- c("Capture1_K4.hist", "Capture2_K4.hist", "Capture3_K4.hist", "Capture4_K4.hist")

for (file in files) {
  print(file)
  files_list <- list.files(pattern=file)
  print(files_list)

  files <- c(paste0(file, ".AllCDS.all.split.txt"), paste0(file, ".AllExon.all.split.txt"), paste0(file
  print(files)

  labs <- c("CDS","Exon","Gene","UTR")

  cov <- list()
  for (i in 1:length(files)) {
    cov[[i]] <- read.table(files[i])[,c(2,5)]
    cov_cumul=1-cumsum(cov[[i]][,2])
    cov[[i]]$cov_cumul <- c(1,cov_cumul[-length(cov_cumul)])
    cov[[i]]$sample=labs[i]
```

```
    }

    cov_df=do.call("rbind",cov)
    names(cov_df)[1:2]=c("depth","fraction")

    pcbPalette <- c("#009E73" ,"#D55E00","#CC79A7","#56B4E9")

    p1 <- ggplot(cov_df, aes(x= depth, y=cov_cumul, color=sample))  + xlim(0,25)+
      scale_alpha(guide = 'none') +
      geom_line(size=1.5)+
      #geom_segment(aes(x=20, y=0, xend=20, yend=1, color="red"))+
      scale_color_manual(values=pcbPalette) +
      scale_fill_manual(values=pcbPalette) +
      ggtitle(file)+
      ylab("% of Bases > Depth")+
      xlab("Depth")+
      theme_bw() +
      theme(plot.title = element_text(size = 12, face = "bold",hjust = 0.5)) +
      theme(legend.title = element_blank()) +
      theme(legend.position=c(0.50,0.75))

    plot_list[[length(plot_list)+1]] <- p1

    png(filename=paste0("Figure2_", file, ".png"), type="cairo",units="px", width=5600,
        height=3000, res=600, bg="transparent")

    print(p1)

    dev.off()

}
```

```
## [1] "Capture1_K4.hist"
## [1] "Capture1_K4.hist.AllCDS.all.split.txt"
## [2] "Capture1_K4.hist.AllExon.all.split.txt"
## [3] "Capture1_K4.hist.AllGene.all.split.txt"
## [4] "Capture1_K4.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture1_K4.hist.png"
## [1] "Capture1_K4.hist.AllCDS.all.split.txt"
## [2] "Capture1_K4.hist.AllExon.all.split.txt"
## [3] "Capture1_K4.hist.AllGene.all.split.txt"
## [4] "Capture1_K4.hist.AllUTR.all.split.txt"

## Warning: Removed 4343 row(s) containing missing values (geom_path).

## [1] "Capture2_K4.hist"
## [1] "Capture2_K4.hist.AllCDS.all.split.txt"
## [2] "Capture2_K4.hist.AllExon.all.split.txt"
## [3] "Capture2_K4.hist.AllGene.all.split.txt"
## [4] "Capture2_K4.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture2_K4.hist.png"
## [1] "Capture2_K4.hist.AllCDS.all.split.txt"
## [2] "Capture2_K4.hist.AllExon.all.split.txt"
## [3] "Capture2_K4.hist.AllGene.all.split.txt"
## [4] "Capture2_K4.hist.AllUTR.all.split.txt"
```

```
## Warning: Removed 5878 row(s) containing missing values (geom_path).

## [1] "Capture3_K4.hist"
## [1] "Capture3_K4.hist.AllCDS.all.split.txt"
## [2] "Capture3_K4.hist.AllExon.all.split.txt"
## [3] "Capture3_K4.hist.AllGene.all.split.txt"
## [4] "Capture3_K4.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture3_K4.hist.png"
## [1] "Capture3_K4.hist.AllCDS.all.split.txt"
## [2] "Capture3_K4.hist.AllExon.all.split.txt"
## [3] "Capture3_K4.hist.AllGene.all.split.txt"
## [4] "Capture3_K4.hist.AllUTR.all.split.txt"

## Warning: Removed 6236 row(s) containing missing values (geom_path).

## [1] "Capture4_K4.hist"
## [1] "Capture4_K4.hist.AllCDS.all.split.txt"
## [2] "Capture4_K4.hist.AllExon.all.split.txt"
## [3] "Capture4_K4.hist.AllGene.all.split.txt"
## [4] "Capture4_K4.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture4_K4.hist.png"
## [1] "Capture4_K4.hist.AllCDS.all.split.txt"
## [2] "Capture4_K4.hist.AllExon.all.split.txt"
## [3] "Capture4_K4.hist.AllGene.all.split.txt"
## [4] "Capture4_K4.hist.AllUTR.all.split.txt"

## Warning: Removed 7587 row(s) containing missing values (geom_path).
```
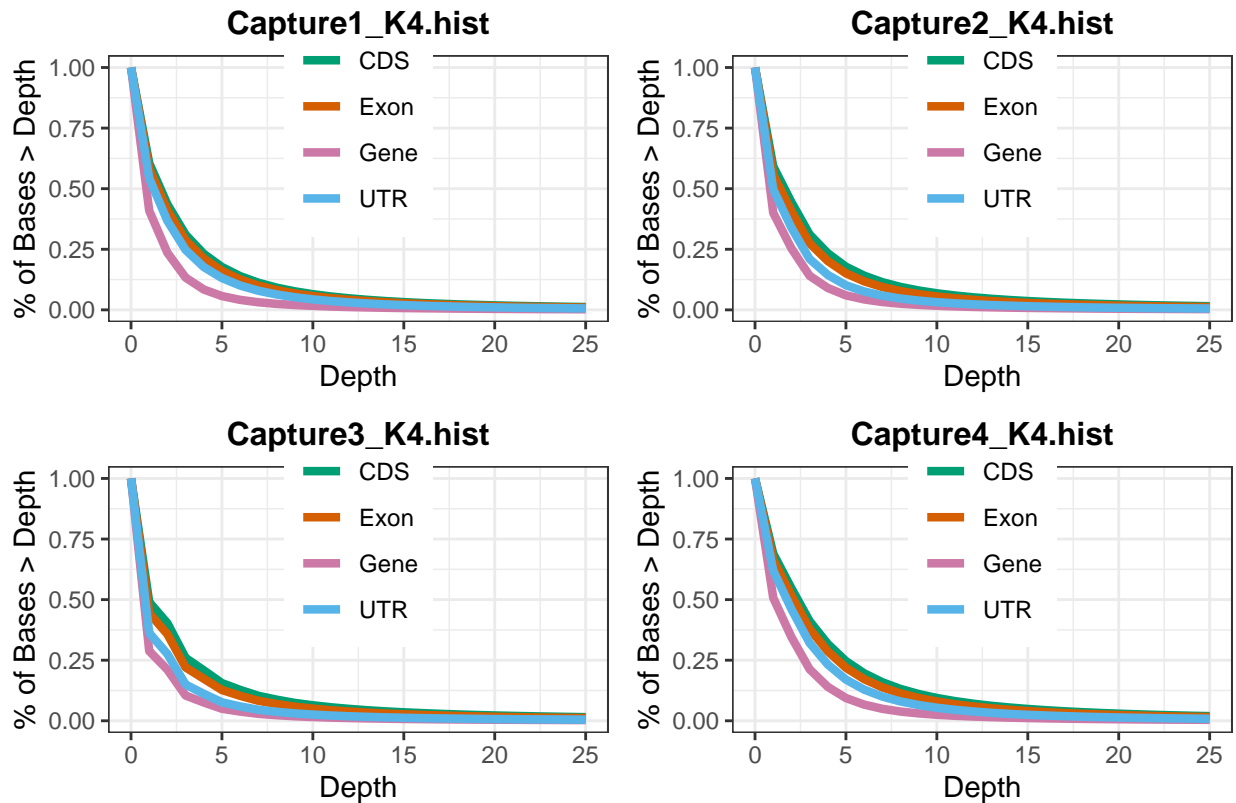
```r
# combine all the ggplot objects in the list using grid.arrange()
pcombined <- grid.arrange(grobs = plot_list, ncol = 2, top=textGrob("Capture B4 Read Depth across genom
```

```
## Warning: Removed 4343 row(s) containing missing values (geom_path).

## Warning: Removed 5878 row(s) containing missing values (geom_path).

## Warning: Removed 6236 row(s) containing missing values (geom_path).

## Warning: Removed 7587 row(s) containing missing values (geom_path).
```

Capture B4 Read Depth across genome


```
#save the combined plot as PNG file
ggsave(file="Figure2_allcapture_K4.png", pcombined)
```

```
## Saving 6.5 x 4.5 in image
```

```
setwd("/home/jgreen/EAGER_OBJ1a/04_coverage_analysis/01_genome_region/")
# initialize an empty list to store ggplot objects
plot_list <- list()

files <- c("Capture1_M3.hist", "Capture2_M3.hist", "Capture3_M3.hist", "Capture4_M3.hist")

for (file in files) {
  print(file)
  files_list <- list.files(pattern=file)
  print(files_list)

  files <- c(paste0(file, ".AllCDS.all.split.txt"), paste0(file, ".AllExon.all.split.txt"), paste0(file
  print(files)

  labs <- c("CDS","Exon","Gene","UTR")

  cov <- list()
  for (i in 1:length(files)) {
    cov[[i]] <- read.table(files[i])[,c(2,5)]
    cov_cumul=1-cumsum(cov[[i]][,2])
    cov[[i]]$cov_cumul <- c(1,cov_cumul[-length(cov_cumul)])
    cov[[i]]$sample=labs[i]
```

```
  }

  cov_df=do.call("rbind",cov)
  names(cov_df)[1:2]=c("depth","fraction")

  pcbPalette <- c("#009E73" ,"#D55E00","#CC79A7","#56B4E9")

  p1 <- ggplot(cov_df, aes(x= depth, y=cov_cumul, color=sample))  + xlim(0,25)+
    scale_alpha(guide = 'none') +
    geom_line(size=1.5)+
    #geom_segment(aes(x=20, y=0, xend=20, yend=1, color="red"))+
    scale_color_manual(values=pcbPalette) +
    scale_fill_manual(values=pcbPalette) +
    ggtitle(file)+
    ylab("% of Bases > Depth")+
    xlab("Depth")+
    theme_bw() +
    theme(plot.title = element_text(size = 12, face = "bold",hjust = 0.5)) +
    theme(legend.title = element_blank()) +
    theme(legend.position=c(0.50,0.75))

  plot_list[[length(plot_list)+1]] <- p1

  png(filename=paste0("Figure2_", file, ".png"), type="cairo",units="px", width=5600,
      height=3000, res=600, bg="transparent")

  print(p1)

  dev.off()

}
```

```
## [1] "Capture1_M3.hist"
## [1] "Capture1_M3.hist.AllCDS.all.split.txt"
## [2] "Capture1_M3.hist.AllExon.all.split.txt"
## [3] "Capture1_M3.hist.AllGene.all.split.txt"
## [4] "Capture1_M3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture1_M3.hist.png"
## [1] "Capture1_M3.hist.AllCDS.all.split.txt"
## [2] "Capture1_M3.hist.AllExon.all.split.txt"
## [3] "Capture1_M3.hist.AllGene.all.split.txt"
## [4] "Capture1_M3.hist.AllUTR.all.split.txt"

## Warning: Removed 7636 row(s) containing missing values (geom_path).

## [1] "Capture2_M3.hist"
## [1] "Capture2_M3.hist.AllCDS.all.split.txt"
## [2] "Capture2_M3.hist.AllExon.all.split.txt"
## [3] "Capture2_M3.hist.AllGene.all.split.txt"
## [4] "Capture2_M3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture2_M3.hist.png"
## [1] "Capture2_M3.hist.AllCDS.all.split.txt"
## [2] "Capture2_M3.hist.AllExon.all.split.txt"
## [3] "Capture2_M3.hist.AllGene.all.split.txt"
## [4] "Capture2_M3.hist.AllUTR.all.split.txt"
```

```
## Warning: Removed 5856 row(s) containing missing values (geom_path).

## [1] "Capture3_M3.hist"
## [1] "Capture3_M3.hist.AllCDS.all.split.txt"
## [2] "Capture3_M3.hist.AllExon.all.split.txt"
## [3] "Capture3_M3.hist.AllGene.all.split.txt"
## [4] "Capture3_M3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture3_M3.hist.png"
## [1] "Capture3_M3.hist.AllCDS.all.split.txt"
## [2] "Capture3_M3.hist.AllExon.all.split.txt"
## [3] "Capture3_M3.hist.AllGene.all.split.txt"
## [4] "Capture3_M3.hist.AllUTR.all.split.txt"

## Warning: Removed 6814 row(s) containing missing values (geom_path).

## [1] "Capture4_M3.hist"
## [1] "Capture4_M3.hist.AllCDS.all.split.txt"
## [2] "Capture4_M3.hist.AllExon.all.split.txt"
## [3] "Capture4_M3.hist.AllGene.all.split.txt"
## [4] "Capture4_M3.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture4_M3.hist.png"
## [1] "Capture4_M3.hist.AllCDS.all.split.txt"
## [2] "Capture4_M3.hist.AllExon.all.split.txt"
## [3] "Capture4_M3.hist.AllGene.all.split.txt"
## [4] "Capture4_M3.hist.AllUTR.all.split.txt"

## Warning: Removed 4684 row(s) containing missing values (geom_path).
```
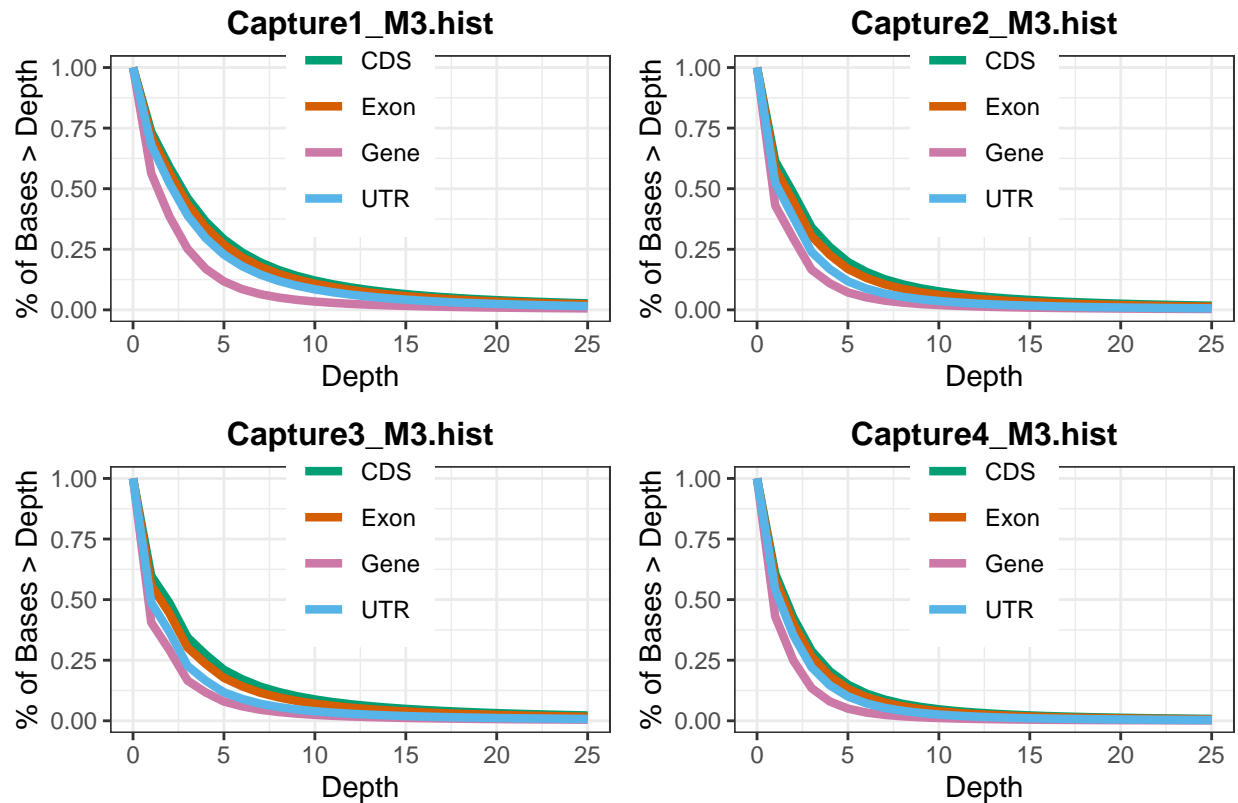
```r
# combine all the ggplot objects in the list using grid.arrange()
pcombined <- grid.arrange(grobs = plot_list, ncol = 2, top=textGrob("Capture B4 Read Depth across genome
```

```
## Warning: Removed 7636 row(s) containing missing values (geom_path).

## Warning: Removed 5856 row(s) containing missing values (geom_path).

## Warning: Removed 6814 row(s) containing missing values (geom_path).

## Warning: Removed 4684 row(s) containing missing values (geom_path).
```

Capture B4 Read Depth across genome


```r
#save the combined plot as PNG file
ggsave(file="Figure2_allcapture_M3.png", pcombined)
```

```
## Saving 6.5 x 4.5 in image
```

```r
setwd("/home/jgreen/EAGER_OBJ1a/04_coverage_analysis/01_genome_region/")
# initialize an empty list to store ggplot objects
plot_list <- list()

files <- c("Capture1_M4.hist", "Capture2_M4.hist", "Capture3_M4.hist", "Capture4_M4.hist")

for (file in files) {
  print(file)
  files_list <- list.files(pattern=file)
  print(files_list)

  files <- c(paste0(file, ".AllCDS.all.split.txt"), paste0(file, ".AllExon.all.split.txt"), paste0(file
  print(files)

  labs <- c("CDS","Exon","Gene","UTR")

  cov <- list()
  for (i in 1:length(files)) {
    cov[[i]] <- read.table(files[i])[,c(2,5)]
    cov_cumul=1-cumsum(cov[[i]][,2])
    cov[[i]]$cov_cumul <- c(1,cov_cumul[-length(cov_cumul)])
    cov[[i]]$sample=labs[i]
```

```
  }

  cov_df=do.call("rbind",cov)
  names(cov_df)[1:2]=c("depth","fraction")

  pcbPalette <- c("#009E73" ,"#D55E00","#CC79A7","#56B4E9")

  p1 <- ggplot(cov_df, aes(x= depth, y=cov_cumul, color=sample))  + xlim(0,25)+
    scale_alpha(guide = 'none') +
    geom_line(size=1.5)+
    #geom_segment(aes(x=20, y=0, xend=20, yend=1, color="red"))+
    scale_color_manual(values=pcbPalette) +
    scale_fill_manual(values=pcbPalette) +
    ggtitle(file)+
    ylab("% of Bases > Depth")+
    xlab("Depth")+
    theme_bw() +
    theme(plot.title = element_text(size = 12, face = "bold",hjust = 0.5)) +
    theme(legend.title = element_blank()) +
    theme(legend.position=c(0.50,0.75))

  plot_list[[length(plot_list)+1]] <- p1

  png(filename=paste0("Figure2_", file, ".png"), type="cairo",units="px", width=5600,
      height=3000, res=600, bg="transparent")

  print(p1)

  dev.off()

}
```

```
## [1] "Capture1_M4.hist"
## [1] "Capture1_M4.hist.AllCDS.all.split.txt"
## [2] "Capture1_M4.hist.AllExon.all.split.txt"
## [3] "Capture1_M4.hist.AllGene.all.split.txt"
## [4] "Capture1_M4.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture1_M4.hist.png"
## [1] "Capture1_M4.hist.AllCDS.all.split.txt"
## [2] "Capture1_M4.hist.AllExon.all.split.txt"
## [3] "Capture1_M4.hist.AllGene.all.split.txt"
## [4] "Capture1_M4.hist.AllUTR.all.split.txt"

## Warning: Removed 5008 row(s) containing missing values (geom_path).

## [1] "Capture2_M4.hist"
## [1] "Capture2_M4.hist.AllCDS.all.split.txt"
## [2] "Capture2_M4.hist.AllExon.all.split.txt"
## [3] "Capture2_M4.hist.AllGene.all.split.txt"
## [4] "Capture2_M4.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture2_M4.hist.png"
## [1] "Capture2_M4.hist.AllCDS.all.split.txt"
## [2] "Capture2_M4.hist.AllExon.all.split.txt"
## [3] "Capture2_M4.hist.AllGene.all.split.txt"
## [4] "Capture2_M4.hist.AllUTR.all.split.txt"
```

```
## Warning: Removed 4953 row(s) containing missing values (geom_path).

## [1] "Capture3_M4.hist"
## [1] "Capture3_M4.hist.AllCDS.all.split.txt"
## [2] "Capture3_M4.hist.AllExon.all.split.txt"
## [3] "Capture3_M4.hist.AllGene.all.split.txt"
## [4] "Capture3_M4.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture3_M4.hist.png"
## [1] "Capture3_M4.hist.AllCDS.all.split.txt"
## [2] "Capture3_M4.hist.AllExon.all.split.txt"
## [3] "Capture3_M4.hist.AllGene.all.split.txt"
## [4] "Capture3_M4.hist.AllUTR.all.split.txt"

## Warning: Removed 5246 row(s) containing missing values (geom_path).

## [1] "Capture4_M4.hist"
## [1] "Capture4_M4.hist.AllCDS.all.split.txt"
## [2] "Capture4_M4.hist.AllExon.all.split.txt"
## [3] "Capture4_M4.hist.AllGene.all.split.txt"
## [4] "Capture4_M4.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture4_M4.hist.png"
## [1] "Capture4_M4.hist.AllCDS.all.split.txt"
## [2] "Capture4_M4.hist.AllExon.all.split.txt"
## [3] "Capture4_M4.hist.AllGene.all.split.txt"
## [4] "Capture4_M4.hist.AllUTR.all.split.txt"

## Warning: Removed 7296 row(s) containing missing values (geom_path).
```
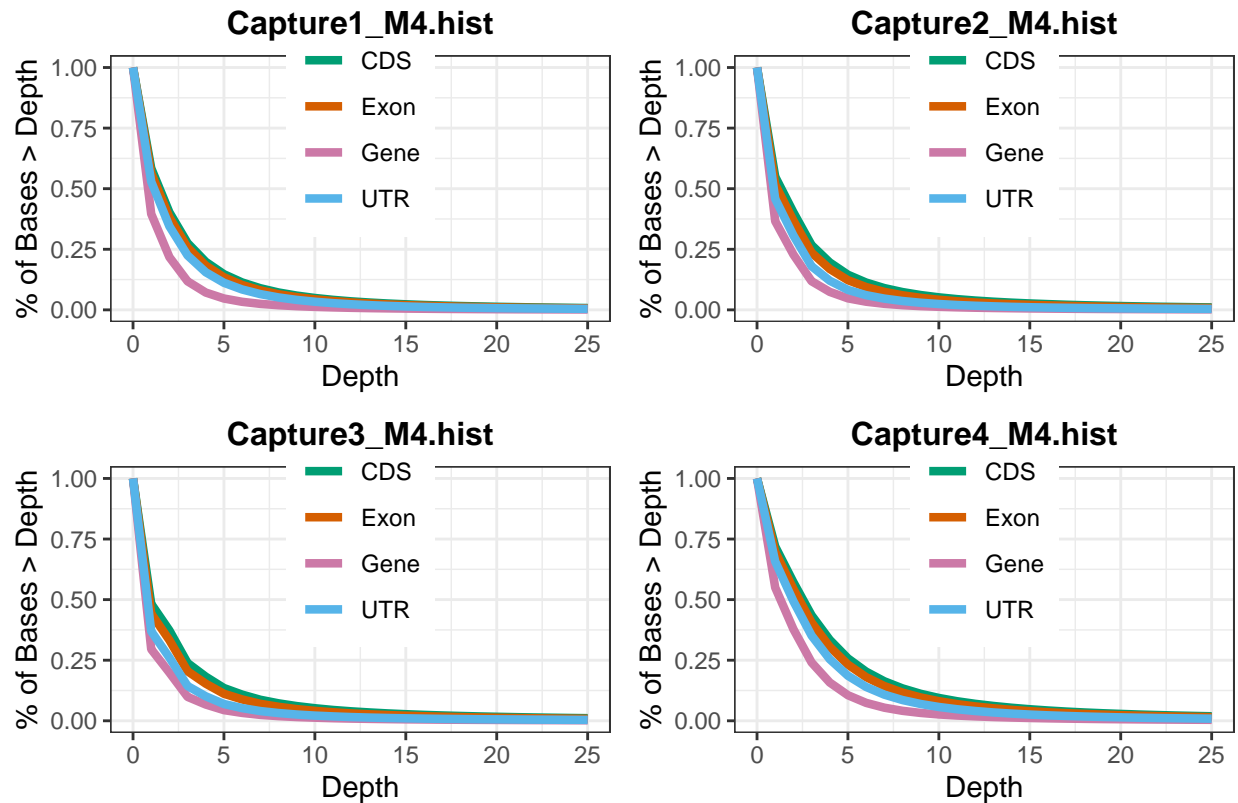
```r
# combine all the ggplot objects in the list using grid.arrange()
pcombined <- grid.arrange(grobs = plot_list, ncol = 2, top=textGrob("Capture B4 Read Depth across genome
```

```
## Warning: Removed 5008 row(s) containing missing values (geom_path).

## Warning: Removed 4953 row(s) containing missing values (geom_path).

## Warning: Removed 5246 row(s) containing missing values (geom_path).

## Warning: Removed 7296 row(s) containing missing values (geom_path).
```

Capture B4 Read Depth across genome


```r
#save the combined plot as PNG file
ggsave(file="Figure2_allcapture_M4.png", pcombined)
```

```
## Saving 6.5 x 4.5 in image
```

```r
setwd("/home/jgreen/EAGER_OBJ1a/04_coverage_analysis/01_genome_region/")
# initialize an empty list to store ggplot objects
plot_list <- list()

files <- c("Capture1_N2.hist", "Capture2_N2.hist", "Capture3_N2.hist", "Capture4_N2.hist")

for (file in files) {
  print(file)
  files_list <- list.files(pattern=file)
  print(files_list)

  files <- c(paste0(file, ".AllCDS.all.split.txt"), paste0(file, ".AllExon.all.split.txt"), paste0(file
  print(files)

  labs <- c("CDS","Exon","Gene","UTR")

  cov <- list()
  for (i in 1:length(files)) {
    cov[[i]] <- read.table(files[i])[,c(2,5)]
    cov_cumul=1-cumsum(cov[[i]][,2])
    cov[[i]]$cov_cumul <- c(1,cov_cumul[-length(cov_cumul)])
    cov[[i]]$sample=labs[i]
```

```
  }

  cov_df=do.call("rbind",cov)
  names(cov_df)[1:2]=c("depth","fraction")

  pcbPalette <- c("#009E73" ,"#D55E00","#CC79A7","#56B4E9")

  p1 <- ggplot(cov_df, aes(x= depth, y=cov_cumul, color=sample))  + xlim(0,25)+
    scale_alpha(guide = 'none') +
    geom_line(size=1.5)+
    #geom_segment(aes(x=20, y=0, xend=20, yend=1, color="red"))+
    scale_color_manual(values=pcbPalette) +
    scale_fill_manual(values=pcbPalette) +
    ggtitle(file)+
    ylab("% of Bases > Depth")+
    xlab("Depth")+
    theme_bw() +
    theme(plot.title = element_text(size = 12, face = "bold",hjust = 0.5)) +
    theme(legend.title = element_blank()) +
    theme(legend.position=c(0.50,0.75))

  plot_list[[length(plot_list)+1]] <- p1

  png(filename=paste0("Figure2_", file, ".png"), type="cairo",units="px", width=5600,
      height=3000, res=600, bg="transparent")

  print(p1)

  dev.off()

}
```

```
## [1] "Capture1_N2.hist"
## [1] "Capture1_N2.hist.AllCDS.all.split.txt"
## [2] "Capture1_N2.hist.AllExon.all.split.txt"
## [3] "Capture1_N2.hist.AllGene.all.split.txt"
## [4] "Capture1_N2.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture1_N2.hist.png"
## [1] "Capture1_N2.hist.AllCDS.all.split.txt"
## [2] "Capture1_N2.hist.AllExon.all.split.txt"
## [3] "Capture1_N2.hist.AllGene.all.split.txt"
## [4] "Capture1_N2.hist.AllUTR.all.split.txt"

## Warning: Removed 6684 row(s) containing missing values (geom_path).

## [1] "Capture2_N2.hist"
## [1] "Capture2_N2.hist.AllCDS.all.split.txt"
## [2] "Capture2_N2.hist.AllExon.all.split.txt"
## [3] "Capture2_N2.hist.AllGene.all.split.txt"
## [4] "Capture2_N2.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture2_N2.hist.png"
## [1] "Capture2_N2.hist.AllCDS.all.split.txt"
## [2] "Capture2_N2.hist.AllExon.all.split.txt"
## [3] "Capture2_N2.hist.AllGene.all.split.txt"
## [4] "Capture2_N2.hist.AllUTR.all.split.txt"
```

```
## Warning: Removed 6967 row(s) containing missing values (geom_path).

## [1] "Capture3_N2.hist"
## [1] "Capture3_N2.hist.AllCDS.all.split.txt"
## [2] "Capture3_N2.hist.AllExon.all.split.txt"
## [3] "Capture3_N2.hist.AllGene.all.split.txt"
## [4] "Capture3_N2.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture3_N2.hist.png"
## [1] "Capture3_N2.hist.AllCDS.all.split.txt"
## [2] "Capture3_N2.hist.AllExon.all.split.txt"
## [3] "Capture3_N2.hist.AllGene.all.split.txt"
## [4] "Capture3_N2.hist.AllUTR.all.split.txt"

## Warning: Removed 5119 row(s) containing missing values (geom_path).

## [1] "Capture4_N2.hist"
## [1] "Capture4_N2.hist.AllCDS.all.split.txt"
## [2] "Capture4_N2.hist.AllExon.all.split.txt"
## [3] "Capture4_N2.hist.AllGene.all.split.txt"
## [4] "Capture4_N2.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture4_N2.hist.png"
## [1] "Capture4_N2.hist.AllCDS.all.split.txt"
## [2] "Capture4_N2.hist.AllExon.all.split.txt"
## [3] "Capture4_N2.hist.AllGene.all.split.txt"
## [4] "Capture4_N2.hist.AllUTR.all.split.txt"

## Warning: Removed 5898 row(s) containing missing values (geom_path).
```
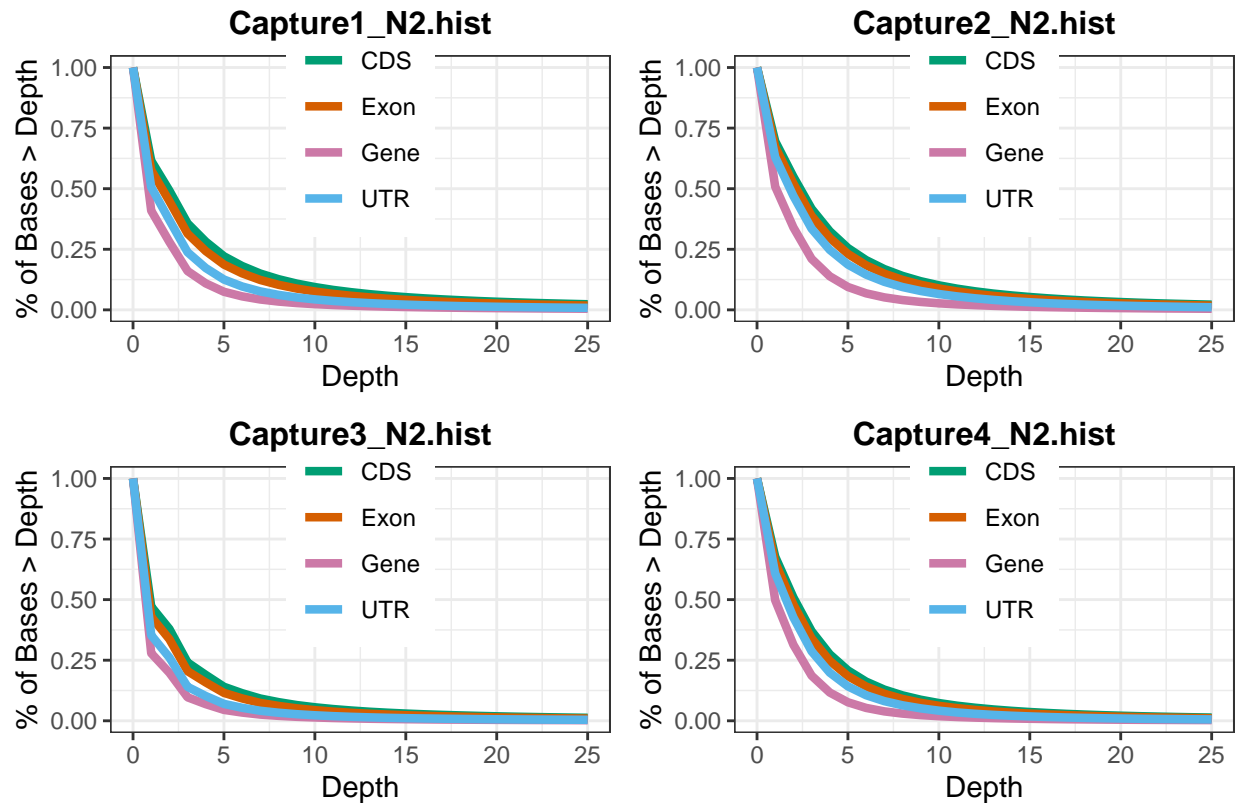
```r
# combine all the ggplot objects in the list using grid.arrange()
pcombined <- grid.arrange(grobs = plot_list, ncol = 2, top=textGrob("Capture B4 Read Depth across genom
```

```
## Warning: Removed 6684 row(s) containing missing values (geom_path).

## Warning: Removed 6967 row(s) containing missing values (geom_path).

## Warning: Removed 5119 row(s) containing missing values (geom_path).

## Warning: Removed 5898 row(s) containing missing values (geom_path).
```

Capture B4 Read Depth across genome

```r
#save the combined plot as PNG file
ggsave(file="Figure2_allcapture_N2.png", pcombined)
```

```
## Saving 6.5 x 4.5 in image
```

```r
setwd("/home/jgreen/EAGER_OBJ1a/04_coverage_analysis/01_genome_region/")
# initialize an empty list to store ggplot objects
plot_list <- list()

files <- c("Capture2_N1.hist", "Capture3_N1.hist", "Capture4_N1.hist")

for (file in files) {
  print(file)
  files_list <- list.files(pattern=file)
  print(files_list)

  files <- c(paste0(file, ".AllCDS.all.split.txt"), paste0(file, ".AllExon.all.split.txt"), paste0(file
  print(files)

  labs <- c("CDS","Exon","Gene","UTR")

  cov <- list()
  for (i in 1:length(files)) {
    cov[[i]] <- read.table(files[i])[,c(2,5)]
    cov_cumul=1-cumsum(cov[[i]][,2])
    cov[[i]]$cov_cumul <- c(1,cov_cumul[-length(cov_cumul)])
    cov[[i]]$sample=labs[i]
```

```
    }

  cov_df=do.call("rbind",cov)
  names(cov_df)[1:2]=c("depth","fraction")

  pcbPalette <- c("#009E73" ,"#D55E00","#CC79A7","#56B4E9")

  p1 <- ggplot(cov_df, aes(x= depth, y=cov_cumul, color=sample))  + xlim(0,25)+
    scale_alpha(guide = 'none') +
    geom_line(size=1.5)+
    #geom_segment(aes(x=20, y=0, xend=20, yend=1, color="red"))+
    scale_color_manual(values=pcbPalette) +
    scale_fill_manual(values=pcbPalette) +
    ggtitle(file)+
    ylab("% of Bases > Depth")+
    xlab("Depth")+
    theme_bw() +
    theme(plot.title = element_text(size = 12, face = "bold",hjust = 0.5)) +
    theme(legend.title = element_blank()) +
    theme(legend.position=c(0.50,0.75))

  plot_list[[length(plot_list)+1]] <- p1

  png(filename=paste0("Figure2_", file, ".png"), type="cairo",units="px", width=5600,
      height=3000, res=600, bg="transparent")

  print(p1)

  dev.off()

}
```

```
## [1] "Capture2_N1.hist"
## [1] "Capture2_N1.hist.AllCDS.all.split.txt"
## [2] "Capture2_N1.hist.AllExon.all.split.txt"
## [3] "Capture2_N1.hist.AllGene.all.split.txt"
## [4] "Capture2_N1.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture2_N1.hist.png"
## [1] "Capture2_N1.hist.AllCDS.all.split.txt"
## [2] "Capture2_N1.hist.AllExon.all.split.txt"
## [3] "Capture2_N1.hist.AllGene.all.split.txt"
## [4] "Capture2_N1.hist.AllUTR.all.split.txt"

## Warning: Removed 11858 row(s) containing missing values (geom_path).

## [1] "Capture3_N1.hist"
## [1] "Capture3_N1.hist.AllCDS.all.split.txt"
## [2] "Capture3_N1.hist.AllExon.all.split.txt"
## [3] "Capture3_N1.hist.AllGene.all.split.txt"
## [4] "Capture3_N1.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture3_N1.hist.png"
## [1] "Capture3_N1.hist.AllCDS.all.split.txt"
## [2] "Capture3_N1.hist.AllExon.all.split.txt"
## [3] "Capture3_N1.hist.AllGene.all.split.txt"
## [4] "Capture3_N1.hist.AllUTR.all.split.txt"
```
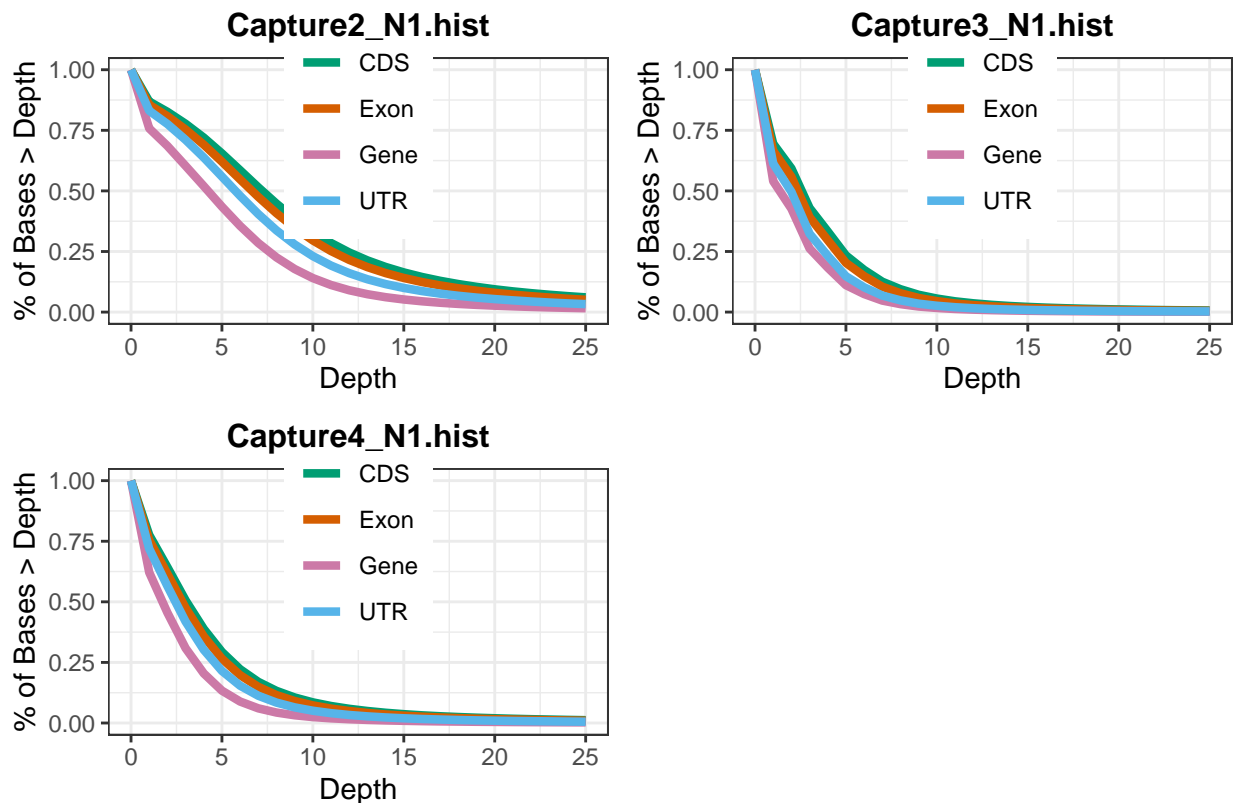
```
## Warning: Removed 4684 row(s) containing missing values (geom_path).

## [1] "Capture4_N1.hist"
## [1] "Capture4_N1.hist.AllCDS.all.split.txt"
## [2] "Capture4_N1.hist.AllExon.all.split.txt"
## [3] "Capture4_N1.hist.AllGene.all.split.txt"
## [4] "Capture4_N1.hist.AllUTR.all.split.txt"
## [5] "Figure2_Capture4_N1.hist.png"
## [1] "Capture4_N1.hist.AllCDS.all.split.txt"
## [2] "Capture4_N1.hist.AllExon.all.split.txt"
## [3] "Capture4_N1.hist.AllGene.all.split.txt"
## [4] "Capture4_N1.hist.AllUTR.all.split.txt"

## Warning: Removed 6621 row(s) containing missing values (geom_path).
```

```r
# combine all the ggplot objects in the list using grid.arrange()
pcombined <- grid.arrange(grobs = plot_list, ncol = 2, top=textGrob("Capture B4 Read Depth across genome
```

```
## Warning: Removed 11858 row(s) containing missing values (geom_path).

## Warning: Removed 4684 row(s) containing missing values (geom_path).

## Warning: Removed 6621 row(s) containing missing values (geom_path).
```


Capture B4 Read Depth across genome

```r
#save the combined plot as PNG file
ggsave(file="Figure2_allcapture_N1.png", pcombined)
```

```
## Saving 6.5 x 4.5 in image
```