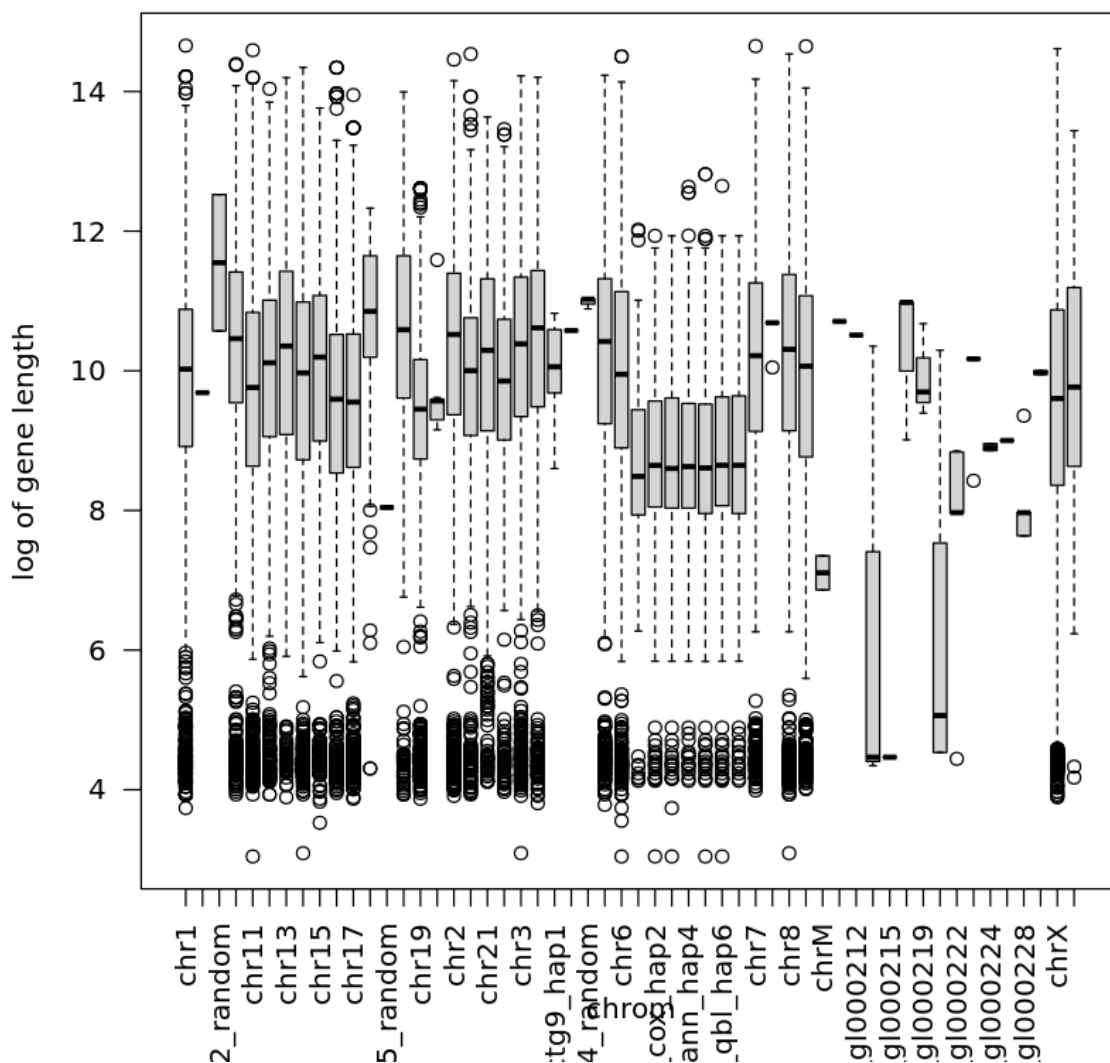


第五讲上机实验

1.用 `read.table` 读入 `he19_gene_table.txt`

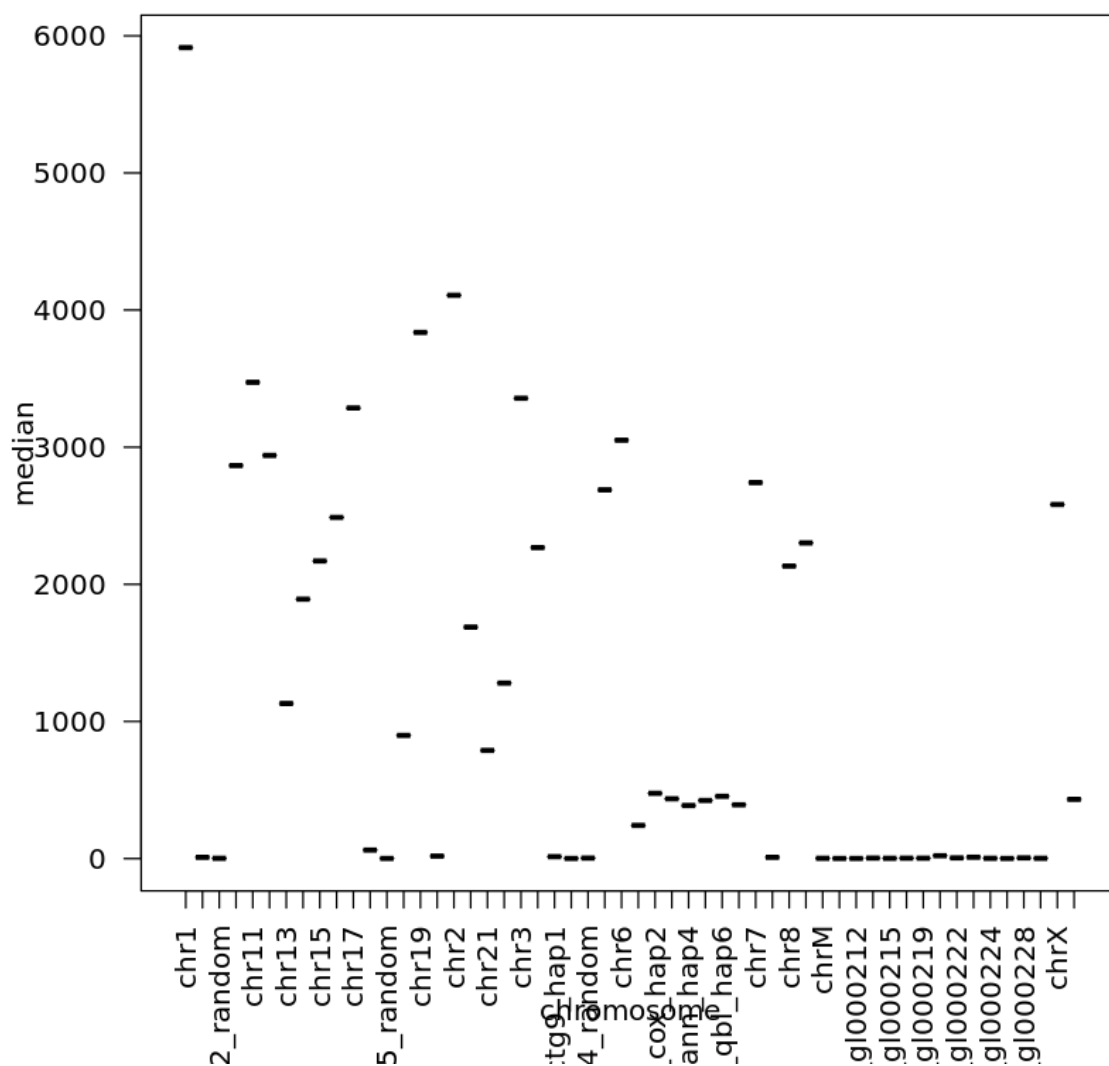
(1) 分染色体画出基因长度取对数后的箱型图

```
1 data <- read.table('/home/ubuntu/R_course/hg19_gene_table.txt', header = T)
2 with(data, boxplot(log(txEnd - txStart + 1) ~ chrom, ylab = 'log of gene
  length', las = 2))
```



(2) 对每条染色体计算基因数量和基因长度的中位数, 分别作为 x 和 y 画出散点图

```
1 data <- read.table('/home/ubuntu/R_course/hg19_gene_table.txt', header = T)
2 gene <- as.data.frame(table(data$chrom))
3 data$geneLength <- data$txEnd - data$txStart + 1
4 lengthMedian <- tapply(data$geneLength, data$chrom, median)
5 plot(as.data.frame(table(data$chrom)), lengthMedian, xlab = 'chromosome',
  ylab = 'median', las = 2)
```



(3) 用函数 `prop.table` 分染色体计算 `+/- strand` 上基因的百分比

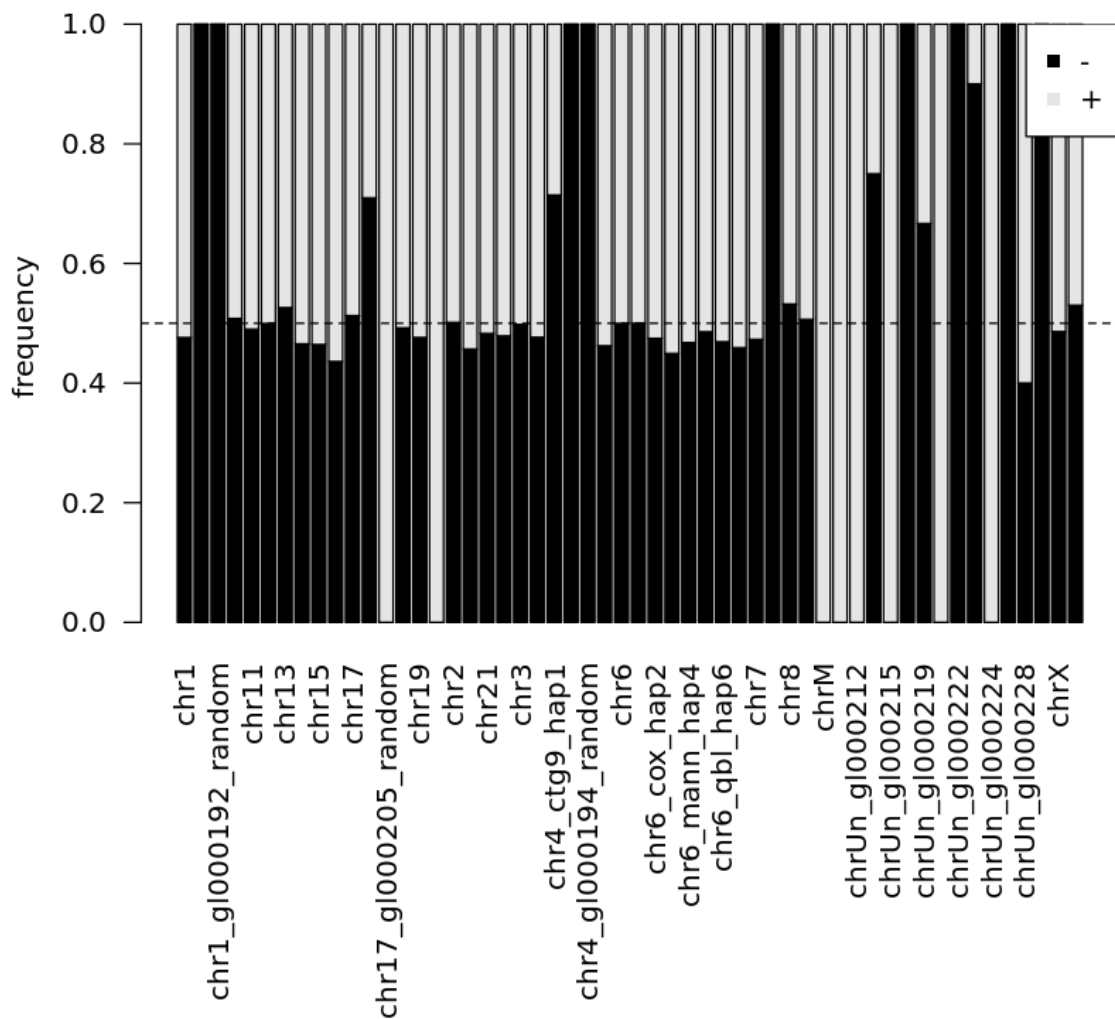
(a) 将上表数值用 `barplot` 画出来

```

1 data <- read.table('/home/ubuntu/R_course/hg19_gene_table.txt', header = T)
2 simData <- as.data.frame(cbind(data$chrom, data$strand))
3 colnames(simData) <- c('chrom', 'strand')
4 # 染色体基因个数
5 table <- prop.table(table(simData$chrom, simData$strand), 1)
6 par(oma=c(8,0,0,0))
7 barplot(t(table), las=2, col=c("black", "grey90"), ylab="frequency", main="gene
8 strand distribution for each chromosome")
9 legend("topright", legend=c("-", "+"), col=c("black", "grey90"), pch=15)
10 abline(h=0.5, lty=2)

```

gene strand distribution for each chromosome



```

1 hm<-read.table('/home/ubuntu/R_course/R_data/gm12878_chr18.txt')
2 hm<-as.matrix(hm)
3 heatmap(hm,Rowv=NA,Colv=NA, labRow=NA,
4         labCol=NA,scale="none",main="figure1")
5 ramp <- colorRamp(c("red", "white"))
6 cols<-rgb( ramp(seq(0, 1, length = 128)), max = 255)
7 heatmap(log(hm+1),col=rev(cols),
8         Rowv=NA,Colv=NA, labRow=NA,
9         labCol=NA,scale="none",main="figure2")
10 heatmap(log(hm[1:100,1:100]+1),col=rev(cols),
11         Rowv=NA,Colv=NA, labRow=NA,
12         labCol=NA,scale="none",main="figure3")
13 ramp <- colorRamp(c("red", "white", "blue"))
14 cols<-rgb( ramp(seq(0, 1, length = 128)), max = 255)
15 heatmap(log(hm[1:100,1:100]+1),col=rev(cols),
16         Rowv=NA,Colv=NA, labRow=NA,
17         labCol=NA,scale="none",main="figure4")

```

figure1

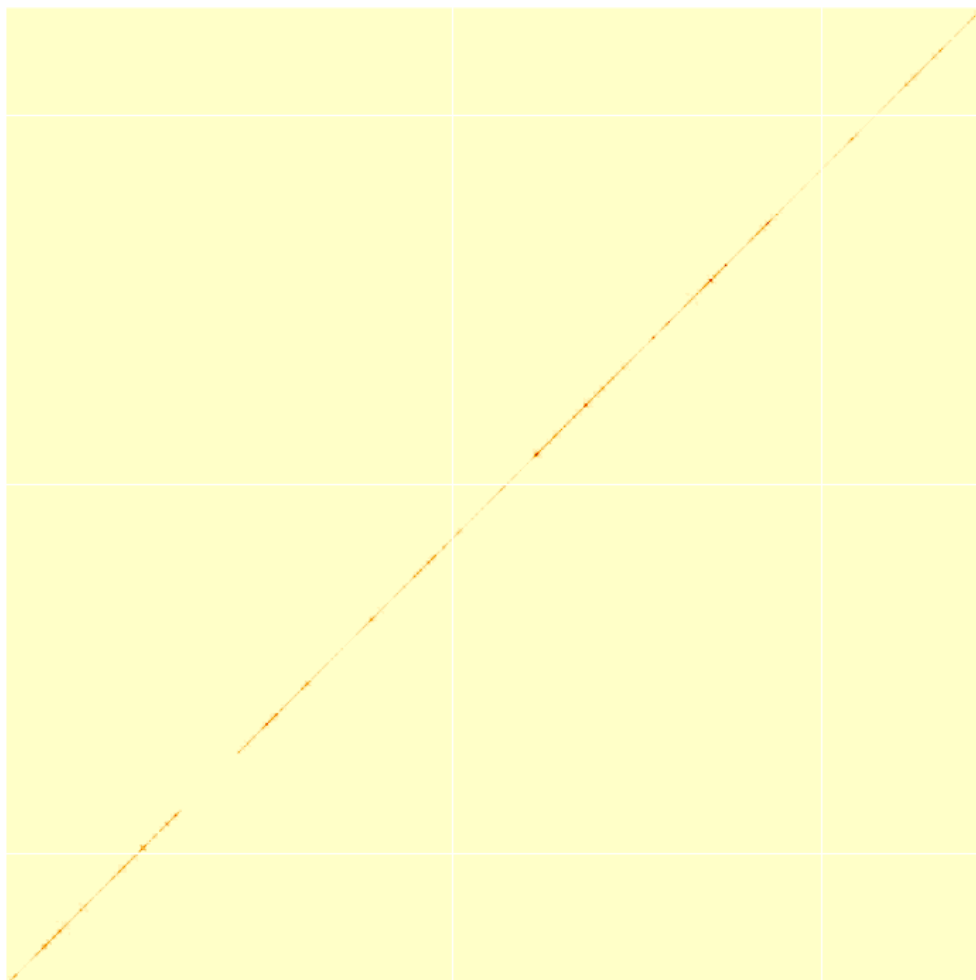


figure2

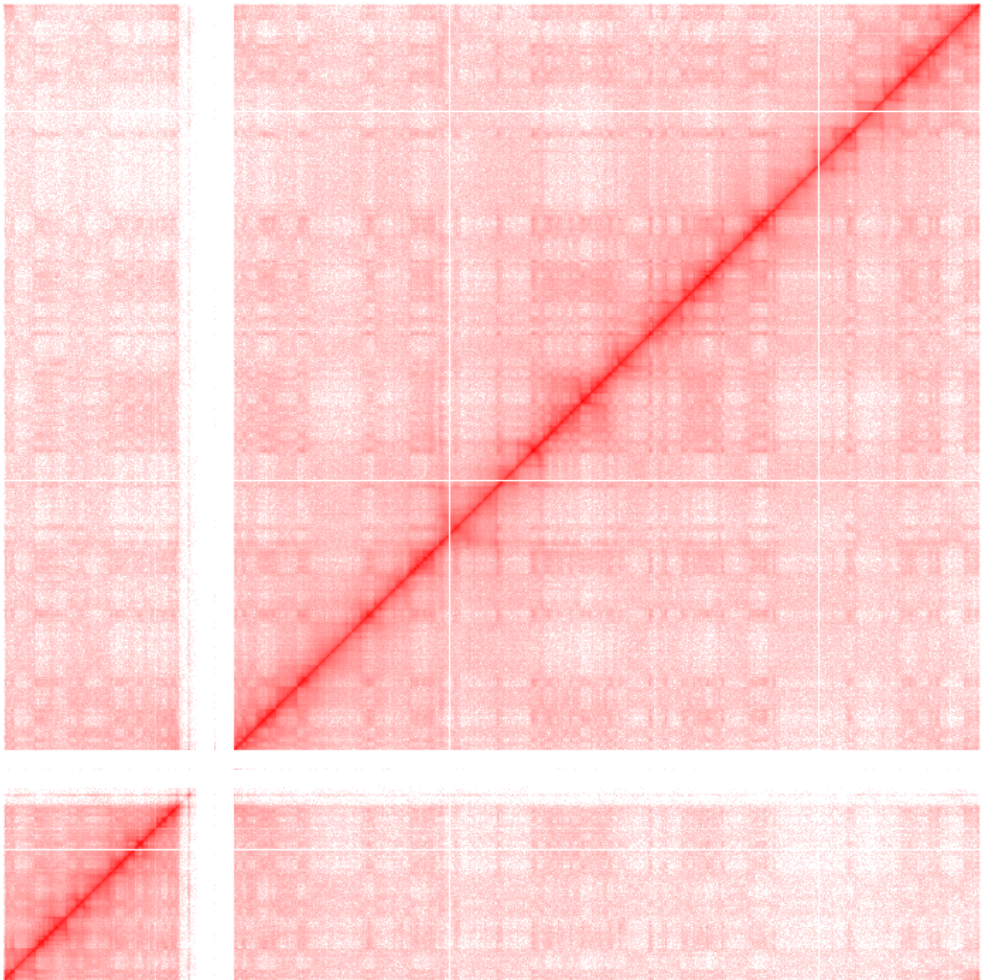


figure3

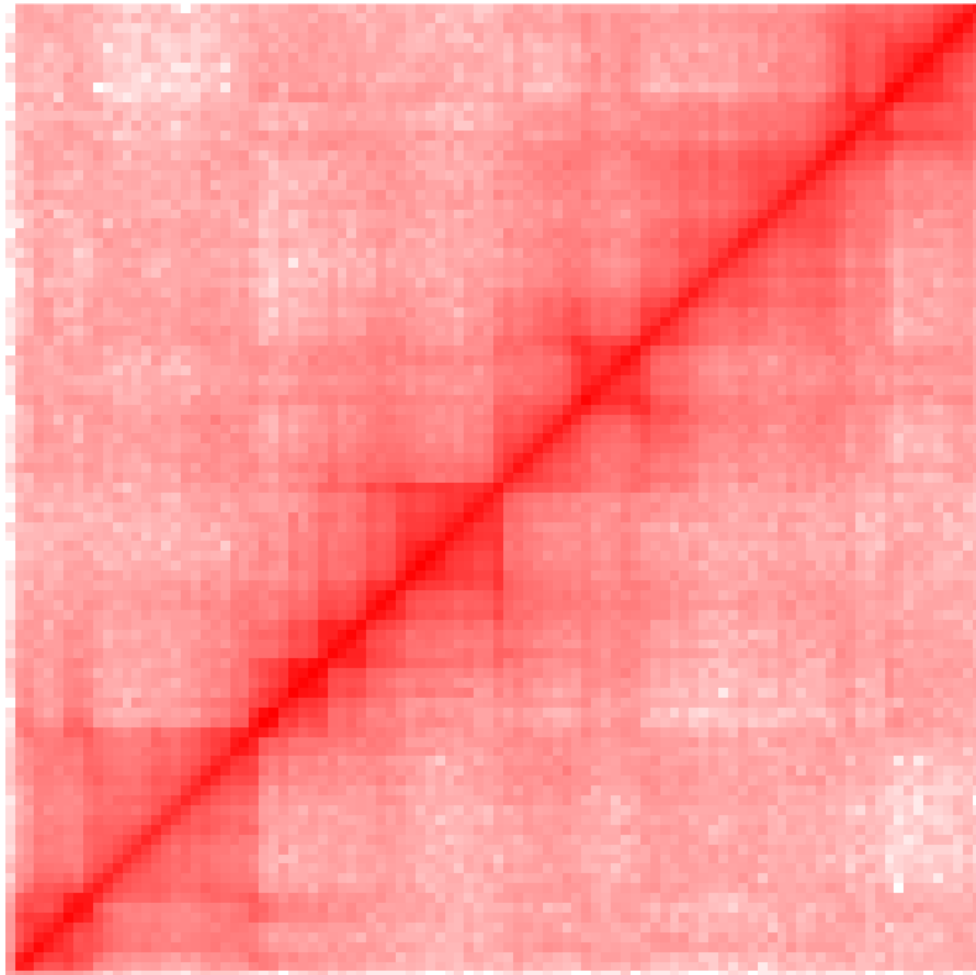
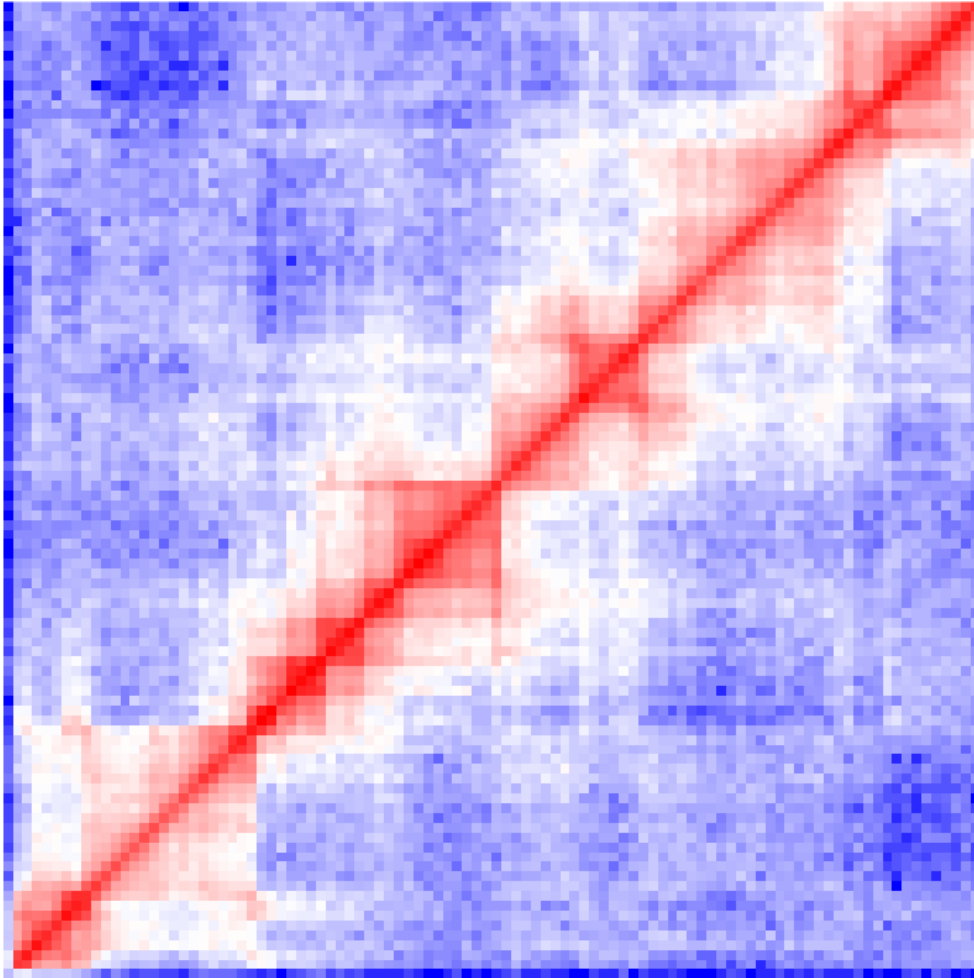


figure4



- 取对数
- 使用不同颜色增加图像对比度