1.读取fasta文件，计算每个基因的GC含量，并按照GC含量从大到小的顺序输出。

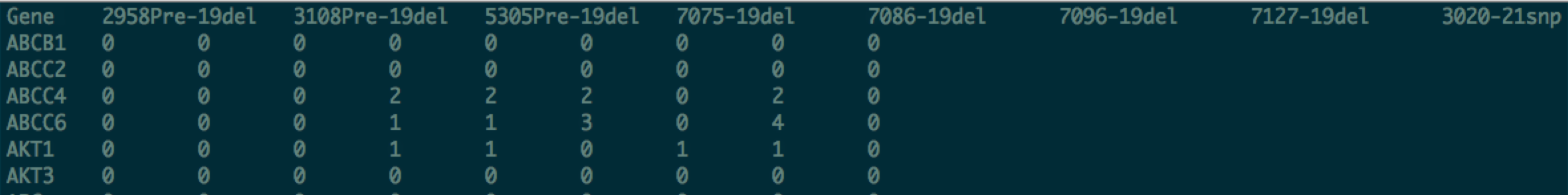
要求：GC含量计算 = （GC碱基个数 / 碱基总数）\* 100,GC含量输出保留两位小数。

输出文件格式：（test01.fasta）

gene2   51.44

gene1   39.88

2. 统计“SNV\_InDel\_data.xls”文件中每个基因在每个样本中的出现次数，输出结果格式如下:



3.根据染色体和位置信息，将不同样本的突变信息进行合并，如果样本没有出现某突变，样本对应的突变信息用“-”表示。

要求；输入文件作为参数

输入文件格式说明：第一列是染色体编号，第二列是位置信息，第三列是参考基因型，第四列是突变信息（突变基因型，突变频率和DP4,分别以“|”隔开），如下：

1 161514631 T C|0.132632|(272,140,47,16)

2 222301193 A C|0.078624|(260,115,23,9)

输出文件格式：

表头：

Chr Position Ref test03\_1 test03\_2 test03\_3

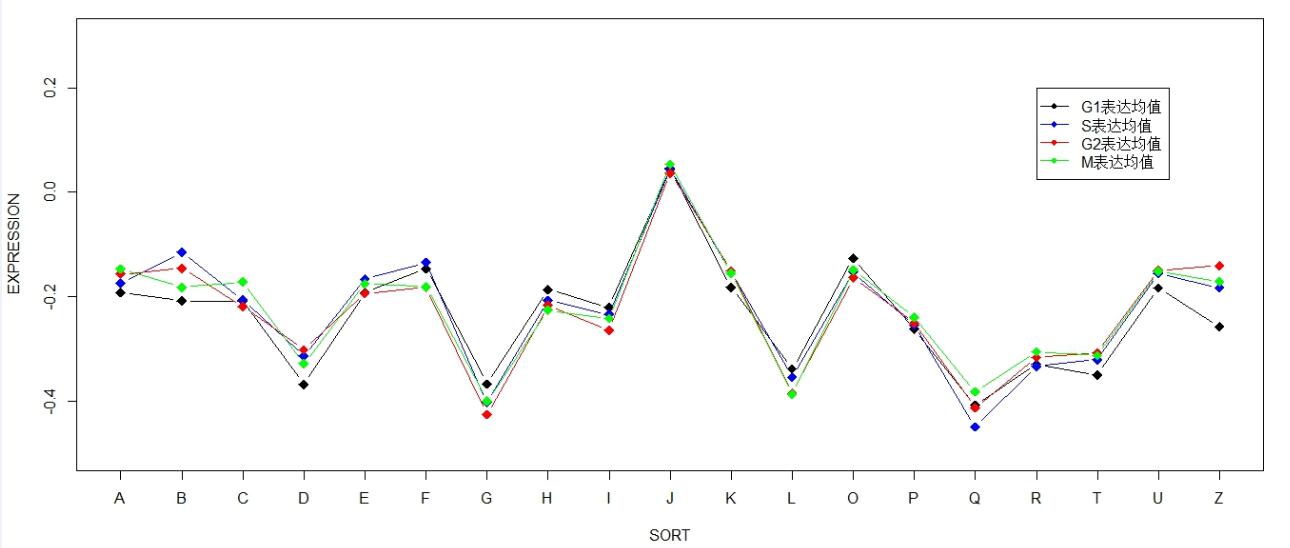
2 222301193 A C|0.181507|(312,166,78,28) C|0.181507|(312,166,78,28) -

20 30671857 C A|0.130909|(336,142,42,30) A|0.130909|(336,142,42,30) -

输入文件路径：test03\_1.info test03\_2.info test03\_3.info

4. 读入文件“01.plot.txt”，第一列为基因功能类别，之后四列为各功能类在G1，S，G2，M时期表达值，绘制折现图，并输出pdf图形文件。要求：显示图例，y轴名称，x轴名称（任意名或参考示例）。

图形示例：



data=read.table("D:/诺禾致源/R/01.plot.txt",sep="\t",header=T)

jpeg("D:/诺禾致源/R/01.plot.jpeg")

x=c(1:dim(data)[1])

y=data[,2]

plot(x,y,type="b",pch=18,col="black",xaxt="n",ylim=c(-1,1),ylab="EXPRESSION",xlab="SORT",cex=1.5)

axis(1,at=(1:19),labels=c("A","B","C","D","E","F","G","H","I","J","K","L","O","P","Q","R","T","U","Z"))

y=data[,3]

lines(y~x,type="b",pch=18,col="blue",cex=1.5)

y=data[,4]

lines(y~x,type="b",pch=18,col="red",cex=1.5)

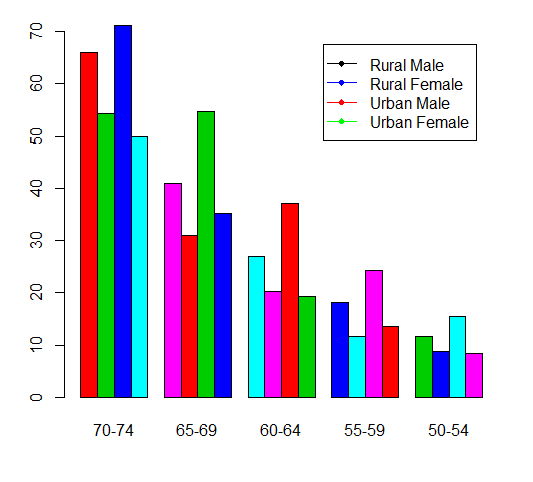
y=data[,5]

lines(y~x,type="b",pch=18,col="green",cex=1.5)

legend("topright",inset=.05,c("G1表达均值","S表达均值","G2表达均值","M表达均值"),col=c("black","blue","red","green"),lty=1,pch=18)

dev.off()

5.读入文件“02.plot.txt”， 弗吉尼亚各年龄段死亡率数据条形图，并输出pdf图形文件。要求：显示标题，y轴名称，x轴名称（任意名或参考示例）。



data=read.table("D:/诺禾致源/R/02.plot.txt",sep="\t",header=T)

pdf("D:/诺禾致源/R/02.plot.pdf")

data=data[,2:6]

GSE <- as.matrix(data)

name=c("70-74","65-69","60-64","55-59","50-54")

colnames(GSE)=name

barplot(GSE,beside=T,col=c(2:6))

legend("topright",inset=.05,c("Rural Male","Rural Female","Urban Male","Urban Female"),col=c("black","blue","red","green"),lty=1,pch=18)

dev.off()