RAProject

Mingzhi Ye

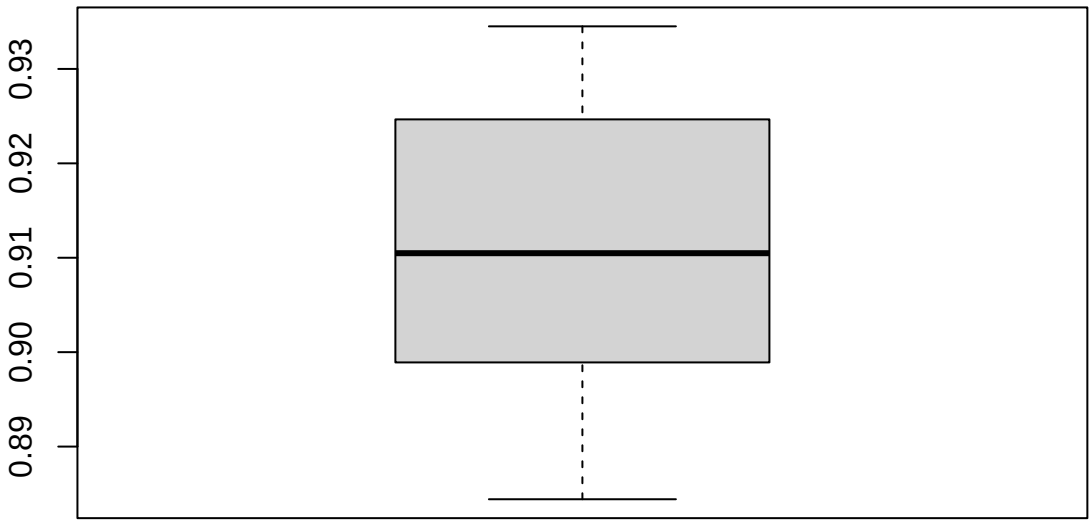
2022/10/10

Calculate distances for samples

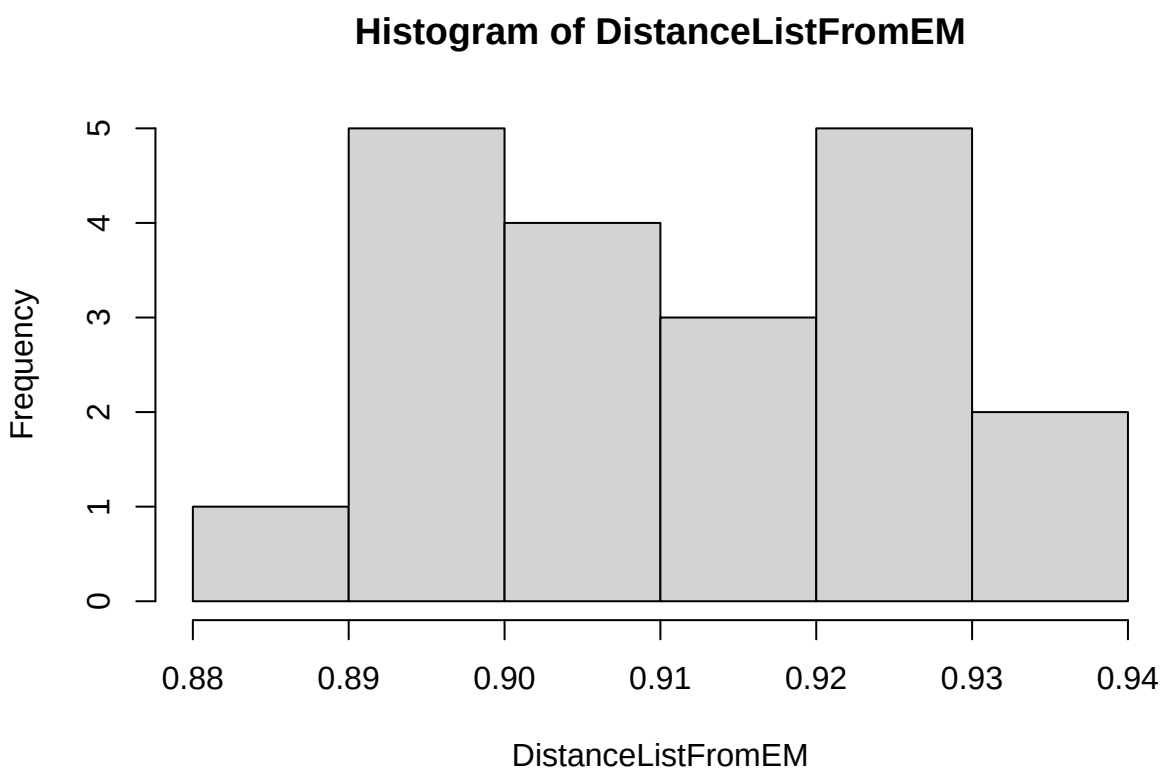
```
DistanceListFromEM=c()
DistanceListFromUN=c()
for(i in 1:20){
  sample=getSample()

  sample$high_confidence<-as.logical(sample$high_confidence)
  sample$productive<-as.logical(sample$productive)
  sample$full_length<-as.logical(sample$full_length)
  sample$is_cell<-as.logical(sample$is_cell)

  samplelist <- split(sample, f = sample$barcode)
  samplelist  <- SplitDataFrameList(samplelist)

  EMpredicted <- clonoStats(samplelist, method = 'EM')
  distance1=getDistance(clonoNames(EMpredicted),clonoAbundance(EMpredicted)[,1])
  DistanceListFromEM[i]=distance1

  UNpredicted <- clonoStats(samplelist, method = 'unique')
  distance2=getDistance(clonoNames(UNpredicted),clonoAbundance(UNpredicted)[,1])
  DistanceListFromUN[i]=distance2
}
```

Visualize

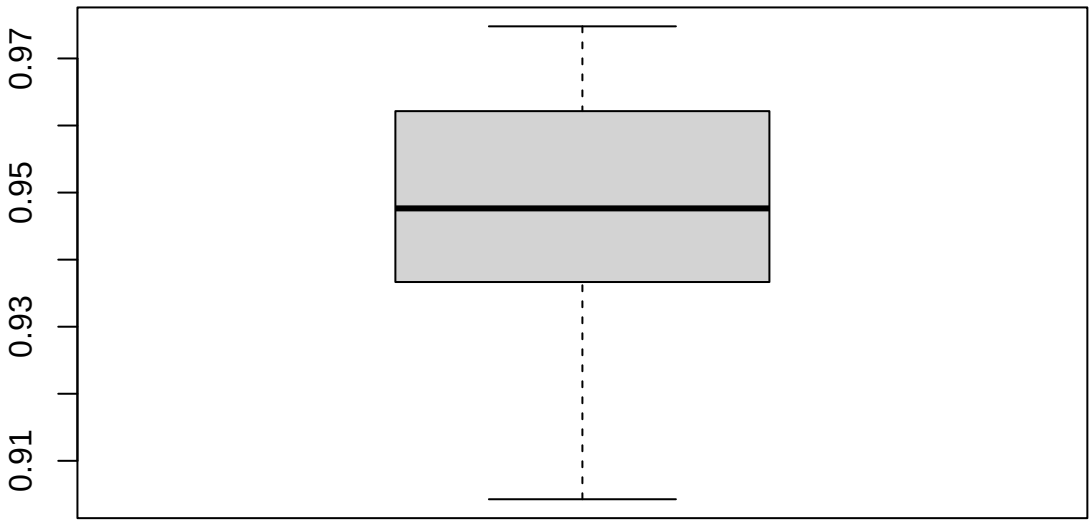
```
boxplot(DistanceListFromEM)
```



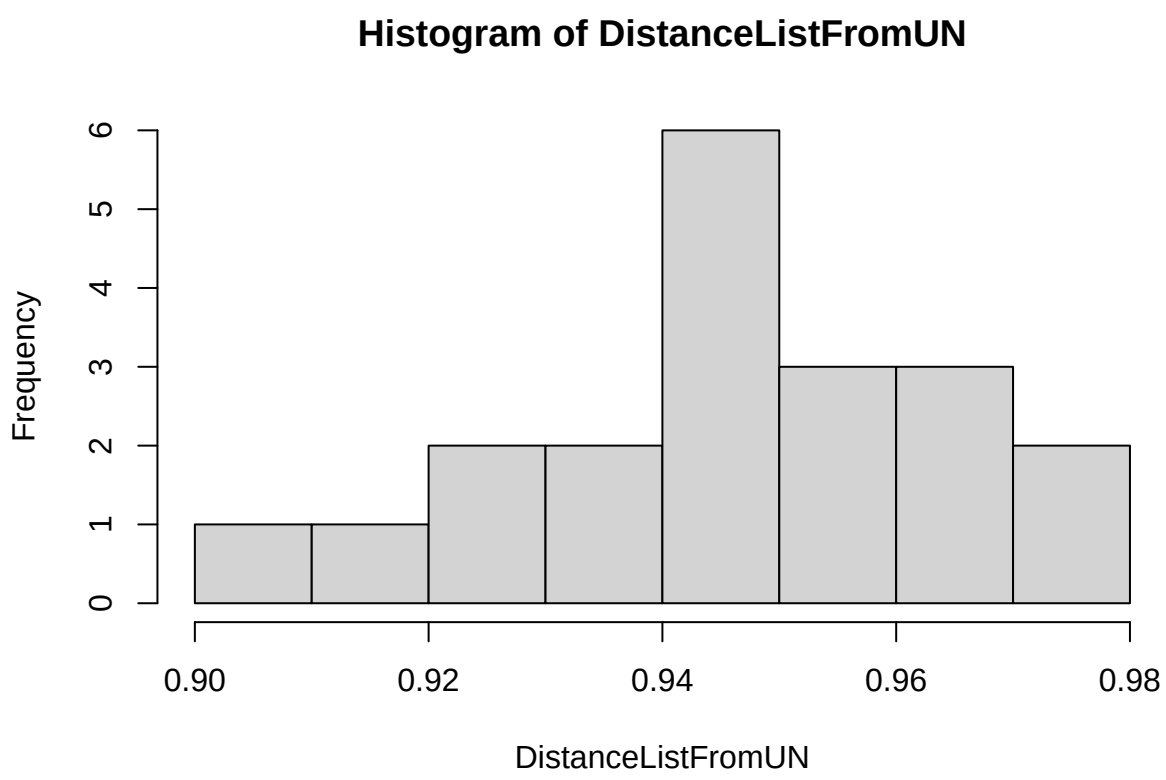

```
hist(DistanceListFromEM)
```




```
boxplot(DistanceListFromUN)
```




```
hist(DistanceListFromUN)
```




```
print(median(DistanceListFromEM))
```

```
## [1] 0.9104828
```

```
print(median(DistanceListFromUN))
```

```
## [1] 0.9476403
```

```
print(mean(DistanceListFromEM))
```

```
## [1] 0.9109828
```

```
print(mean(DistanceListFromUN))
```

```
## [1] 0.9467265
```