Tags used in the Viromics project (metagenomics of viruses in the host-parasitoid community)

In total we had 168 samples (55 CO1, 55 WGA, 55 WTA + 3 negative controls), each tagged with an 8bp tag. See how many differences among them :

```
library(Biostrings)
```

```
## Warning: package 'IRanges' was built under R version 4.3.1
```

```
## Warning: package 'GenomeInfoDb' was built under R version 4.3.1
```

```
tab=read.table("../../Tags.txt", h=T, sep="\t")
myStrings=tab$X8bp_tag_sequence[1:168]

N <- length(myStrings)
myDists <- matrix(0, nrow = N, ncol = N)
for (i in 1:(N-1))
  for (j in (i+1):N)
    myDists[i, j] <- myDists[j, i] <-
  neditStartingAt(myStrings[[i]], myStrings[[j]])
```

```
table=table(myDists[lower.tri(myDists, diag = FALSE)])
barplot(table, xlab="number of differences", ylab = "number of pairwise comparisons", main="")
```
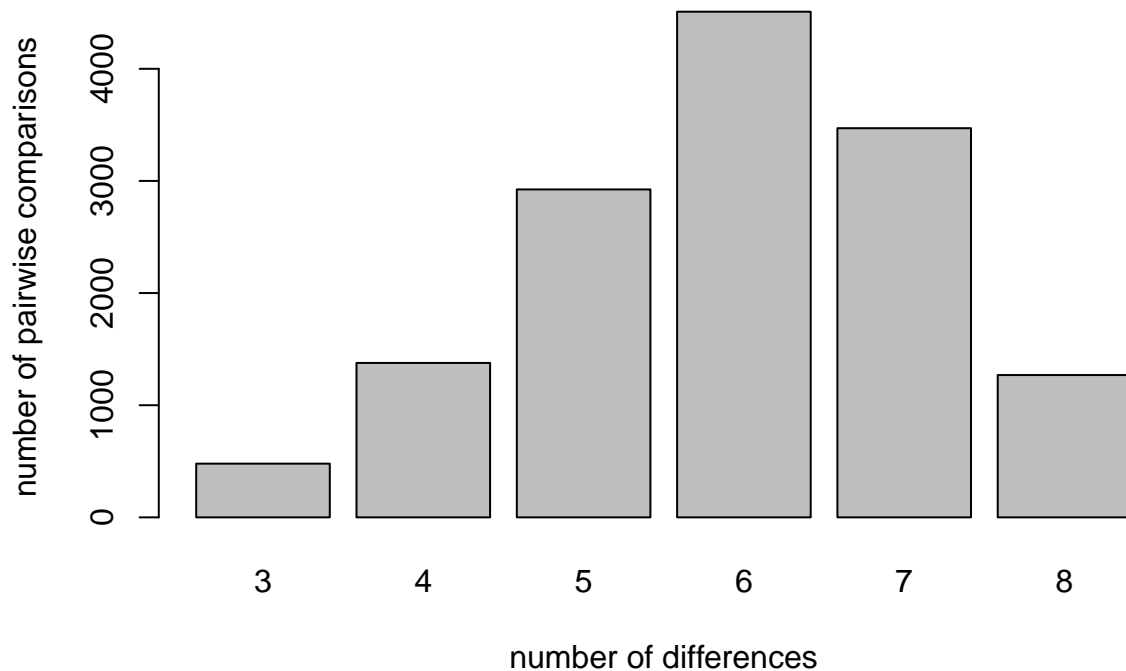

Summary statistics :

```
summary(myDists[lower.tri(myDists, diag = FALSE)])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   3.000   5.000   6.000   5.921   7.000   8.000
```

ref: Meyer, Stenzel, et Hofreiter, « Parallel Tagged Sequencing on the 454 Platform ».10.1038/nprot.2007.520