

# Untitled

*Barry Grant*

*3/2/2018*

## A qucik report

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
library(bio3d)
```

```
## Warning: package 'bio3d' was built under R version 3.4.4
```

Note if you want to run this you will need to install the bio3d package. This would be a one time only thing

```
#install.packages(bio3d)
```

Read your alignment

```
## Read your alignment file
```

```
raw <- read.fasta("muscle+labnanobodies_truncname.fst", rm.dup=FALSE)
```

Clean IDs and write a new alignment file

```
new <- raw
#new$id <- paste0(1:length(raw$id), "_", raw$id)
#write.fasta(new, file="alignment_cleanIDs.fasta")
```

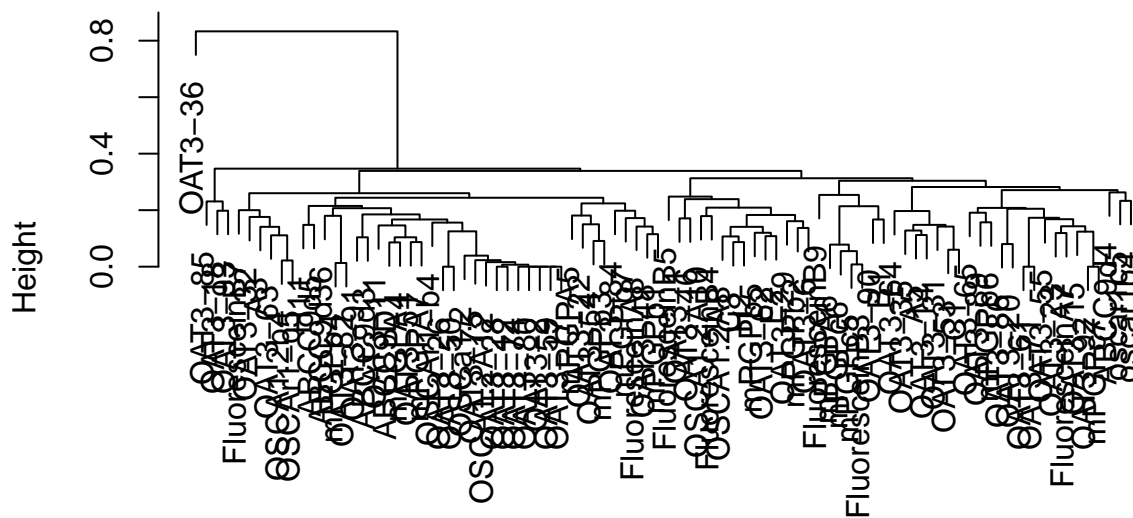
```
## Sequence identity (smiliarity)
```

```
ide <- seqidentity(new, similarity = TRUE)
```

```
hc.all <- hclust(as.dist(1-ide))
```

```
plot(hc.all)
```

## Cluster Dendrogram



```
as.dist(1 - ide)
hclust (*, "complete")
```

Trim to your region and do analysis

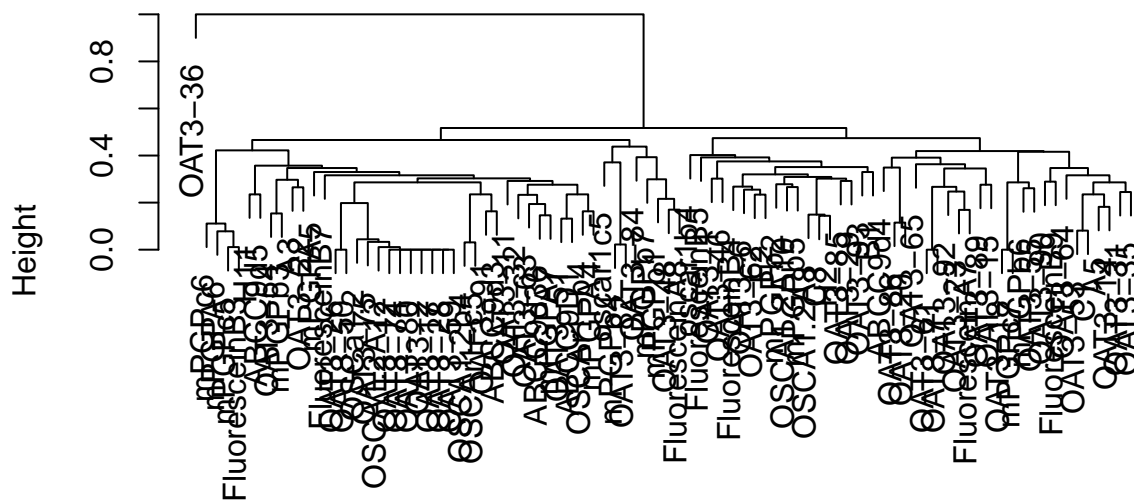
```
## Trim to your region 49 to 157
trim.aln <- new
trim.aln$ali <- trim.aln$ali[,49:157]

## Sequence identity of trimmed alignment section.
trim.ide <- seqidentity(trim.aln)

hc.trim <- hclust(as.dist(1-trim.ide))

plot(hc.trim)
```

## Cluster Dendrogram



```
as.dist(1 - trim.ide)
hclust (*, "complete")
```

Make seperate PDFs

```
pdf("plot_trim_tree.pdf", width=14)
plot(hc.trim)
dev.off()
```

```
## pdf
## 2
```

```
pdf("plot_full_tree.pdf", width=14)
plot(hc.all)
dev.off()
```

```
## pdf
## 2
```