Clustering Gene Expression in East African Cichlids

Madison Gosnell[¥], Sally Sagarsee[§], Peyton Blount[¥], Naureen Aslam[¥], Jeff Kinne[§], Hugo F. Gante[¥]
Indiana State University; [§]Department of Computer Science and [¥]Department of Biology

INTRODUCTION

Coloration differences between distinct species arise from a combination of gene expression and cellular interaction differences. Previous gene expression analyses between two *Neolamprologus* species has suggested thousands of genes are differentially expressed between skin patches. Understanding how these genes function will inform us how the cichlids get their patterns. One way to prioritizing which genes to study is through clustering based on the level of expression they show. Having a control fish (*N. gracilis*) will help us achieve just that.

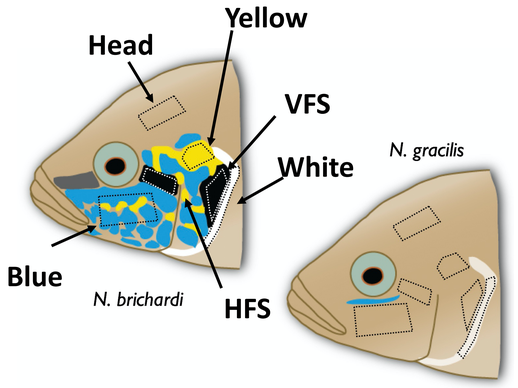

Figure 1. *N. brichardi* are the more "colorful" of the two as *N. gracilis* are our control in this case.

OBJECTIVE

The goal of this project is to prioritize genes differentially expressed in the facial skin of *N. brichardi* and *N. gracilis* fishes for further studies of the effect on the patterns and colors of the fish.

METHODS

To begin clustering we first load the necessary data. The following line shows how to load data for the Horizontal Facial Stripe of HFS:

```
HFS_data <- read.csv("HFS_Black_stripes_DEgenes.txt", sep = "\t")

# function to create fig. 2 and fig. 3
cluster_columns <- function(which_cols, title){
  first_dist <- dist(t(all_data2[which_cols]), method = "euclidean")
  hc_first <- hclust(first_dist, method = "complete")
  plot(hc_first, cex = 0.6, hang = -1, main = title)
}
```

EXPECTED RESULTS

It was first thought that skin patches with similar colors would show closer expression in the genes tested. However it was found that the closer the patches are to each other, the closer the level of expression as seen in **Figures 2 and 3**. In Figure 2. the first cluster is yellow and white. In Figure 3, the first cluster is white, VFS, and yellow. This indicates that position effects and tissue environment guides most of gene expression.

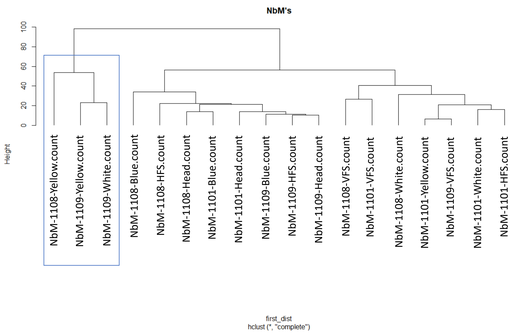

Figure 2. Shows *N. brichardi* males. The first cluster consists of white and yellow.

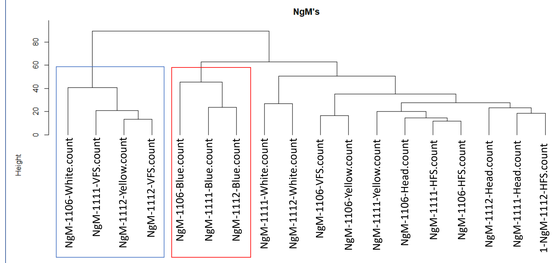

Figure 3."Control" samples that shows *N. gracilis* males. The first cluster consists of VFS, yellow, and white.

DISCUSSION

Clustering the differentially expressed genes in these fish can potentially help prioritizing efforts to study how they influence color pattern differences between species. This approach has been successfully used in previous studies (e.g., Henning *et al.* 2013), where different genes influencing color change cluster according to the type of change observed. Having the control fish (*N. gracilis)* helps to compare the level of expression in the two fishes and gives us a better understanding of the *N. brichardi*.

Ultimately this approach will help select a few of the most promising candidates from a list of several thousands for further functional analyses.

It was interesting to see that there wasn't a defined cluster of one color in the *N. brichardi,* but there was in the *N. gracilis* (highlighted in red in Fig. 3).

REFERENCES

Henning et al.: **Transcriptomics of morphological color change in polychromatic Midas cichlids.** *BMC Genomics 2013 14:171.*

ACKNOWLEDGEMENTS

Rusty Gonser
Naureen Aslam
NIH grant 5R25MD011712