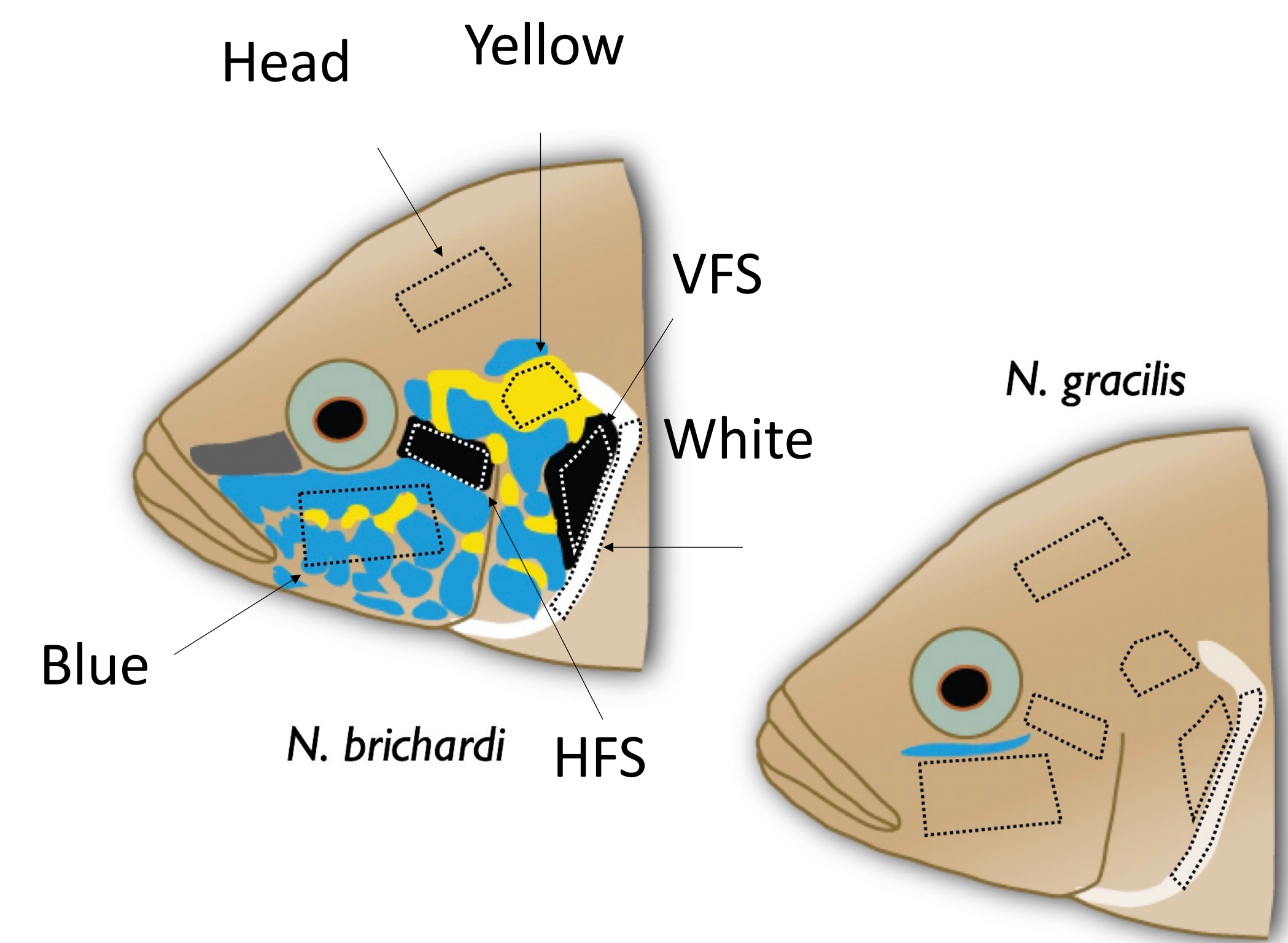# Clustering Gene Expression in East African Cichlids

Madison Gosnell[1], Peyton Blount[2], Jeff Kinne[1], Hugo F. Gante[2]

Indiana State University, [1]Department of Computer Science, [2]Department of Biology

Figure 1. *Neolamprologus brichardi* have colorful facial pigmentation. Facial pattern is absent in *N. gracilis* (used as control).
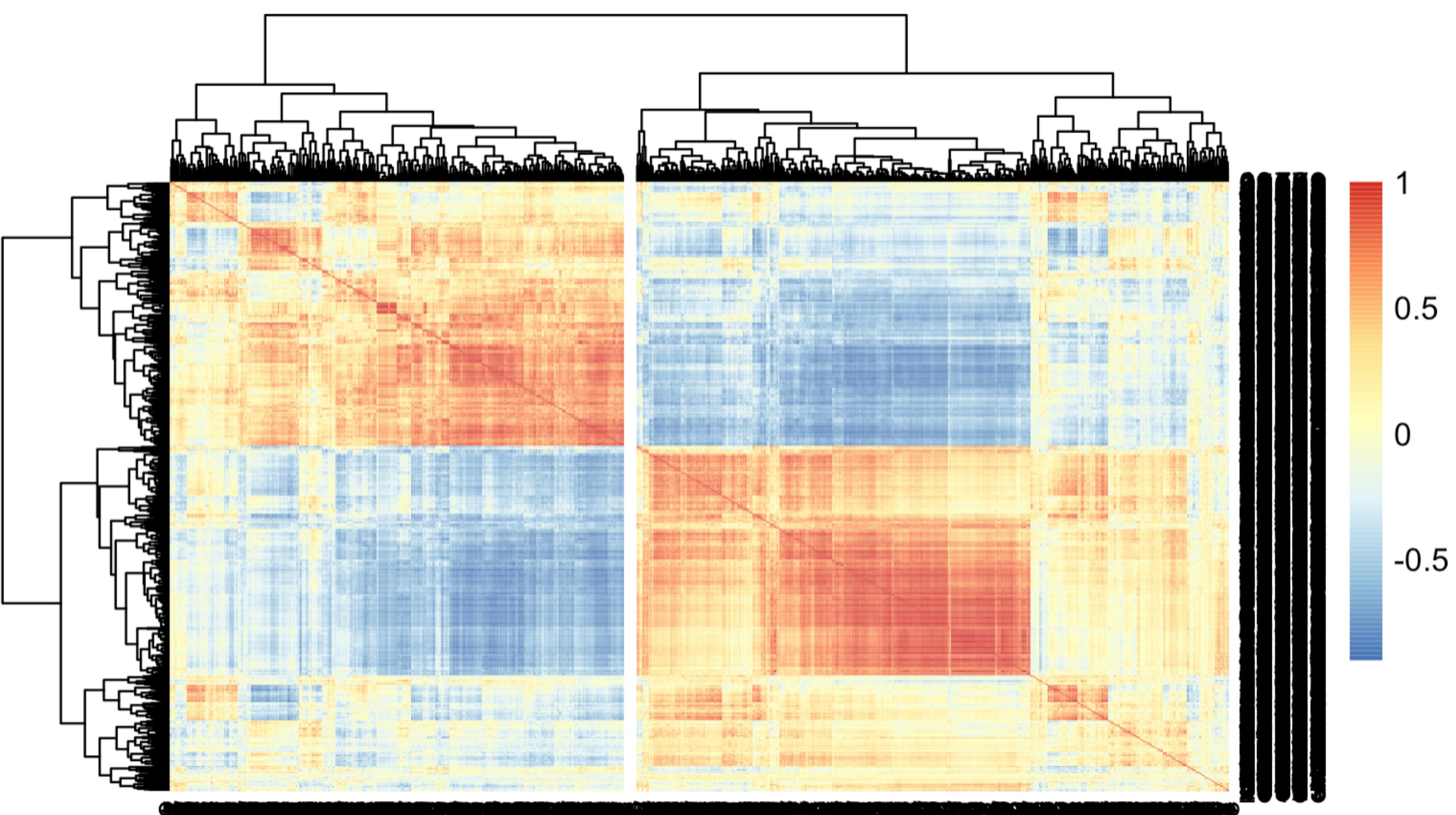
INTRODUCTION

Coloration differences between distinct species arise from a combination of gene expression and cellular interaction differences. Previous gene expression analyses between two *Neolamprologus* species have 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 across color patches, while controlling for position effects. Having a control fish (*N. gracilis*) will help us achieve just that.

OBJECTIVE

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

METHODS

To get a smaller group of genes out of over 38,000 genes expressed in the facial skin, we first created two linear models, one based on 'patch location' and another based on 'patch location' and 'species'. From their comparison, we identified a group of 738 genes that could possibly be causally involved in pigmentation. To visualize groups with similar characteristic, patches and genes were separately clustered (**figs. 2–4**).

Figure 3. Clustering of *N. brichardi* samples. Visible split between the anterior facial patches and the more posterior facial patches.

Figure 4. Clustering of *N. gracilis* samples. Clusters show a less clear anterior-posterior split in gene expression differences.
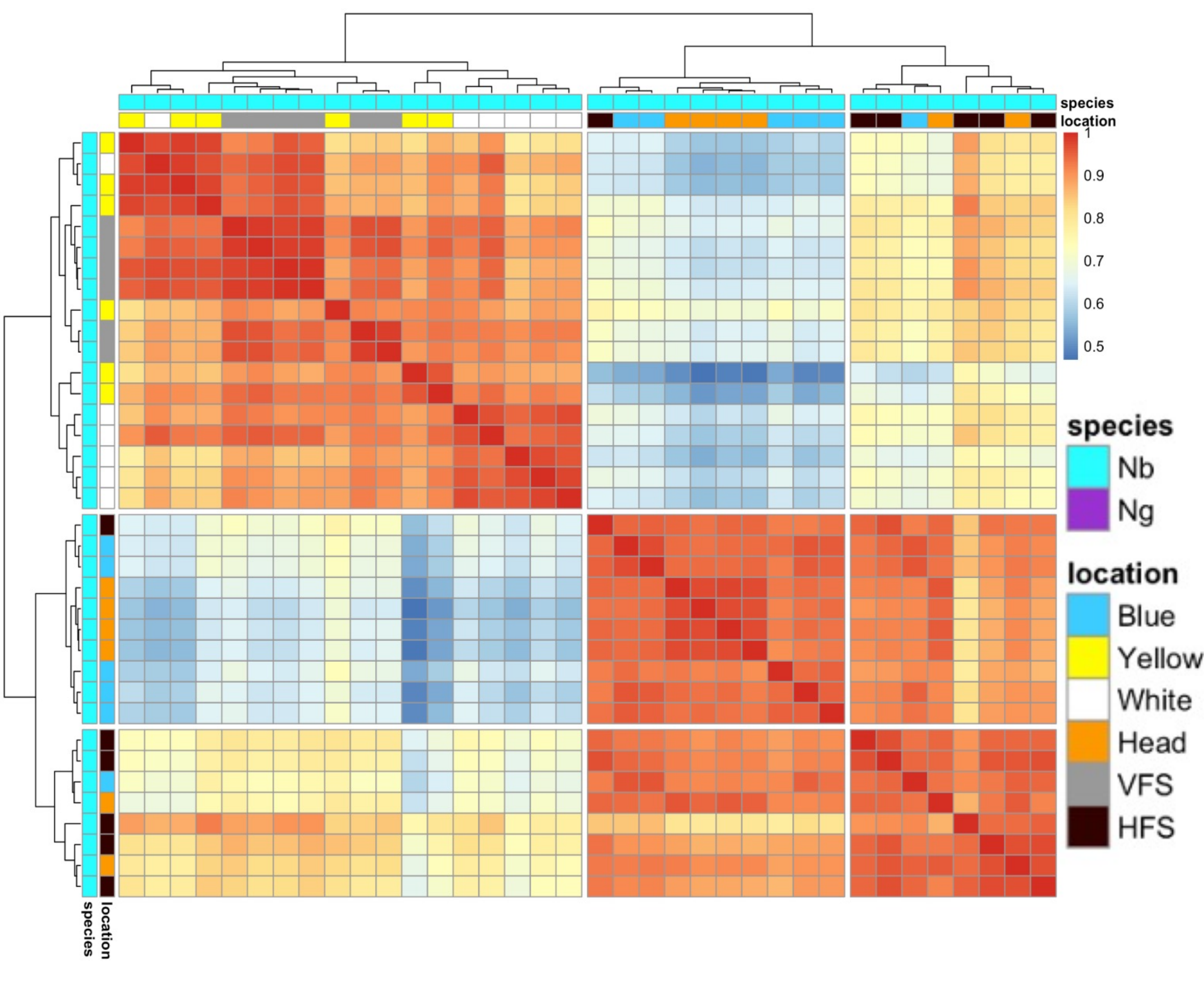

Figure 2. Clustering of 738 genes that shows species-level differences in expression.
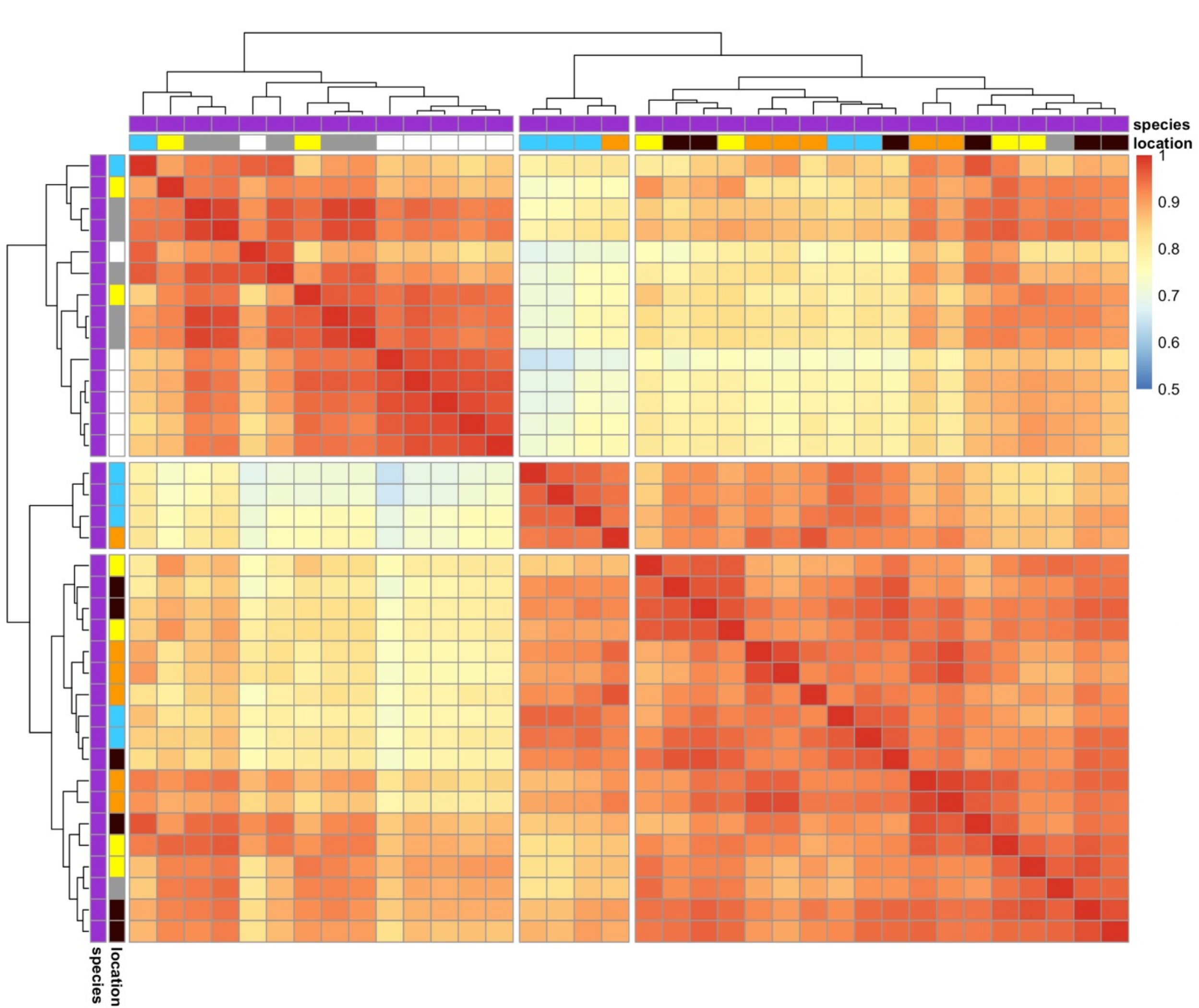
RESULTS and CONCLUSION

There is about the same number of over- and underexpressed genes out of 738 expressed genes differentially between the two species (**fig. 2**). A strong separation between the three anterior patches (head, blue, HFS) from the posterior patches (VFS, yellow, white) is visible in *N. brichardi* but less so in *N. gracilis* (**figs. 3&4**).

Further work remains to analyze the clusters of genes to prioritize genes for further study that are most likely important in pigmentation.