EEB313: Project proposal

Apistogramma is a highly specious genus of cichlid found in habitats across most of South America. These fish live in a broad range of water conditions and vary by species in the stripes and spots that make up the dark portion of their patterns. For this project I will explore the ecological factors that might explain the variation and convergences in *Apistogramma* phenotype. I will examine ecological and phenotypic traits for approximately 100 species in the genus (Römer, 2006). I believe that the visual conditions where the fish live are responsible for the dark patterns and general phenotype of the species that evolve, and that convergence within the genus is due to environment, not shared ancestry. I predict that analysis will cluster species by phenotype best when explained by water type, as this encapsulates numerous ecological factors and determines the visual environment for this highly communicative species.

The data come from two tables published in a book by Dr. Uwe Römer, a lecturer at Universität Trier and prolific writer on the genus *Apistogramma* in scientific literature. The book is a compilation of the author's doctoral and post-doctoral work, with the addition of extensive citation. The first table (Table 2: Habitat preferences and degree of specialisation of *Apistogramma*-species, pp. 194-195) is a compilation of categorical ecological data for 101 species of *Apistogramma*, including water body type (creek, river, general, rock), water type (blackwater, whitewater, clearwater), habitat type (forested, open), and river system (Amazon, Orinoco, Paraguay, other). The second table (Basic data for clusteranalysis of relationships within Apistogramma-species, pp. 1290-1297) is a compilation of 51 characters used for phylogenetic cluster analysis of the same *Apistogramma* species. The phylogenetic tree produced by the author using this table has not been widely accepted, cited, or otherwise deemed successful. These characters include physical traits like colour pattern, body shape, and dentition, as well as reproductive behaviour. The author has given most traits a binary value of 1 or 0, but some traits include values up to 3. I plan to knit the habitat data with the trait data by species to be able to identify ecological factors that might give rise to certain patterns in the character traits, regardless of phylogeny. The habitat data are currently presented as categories that will need to be converted into values. The behavioural traits from the cluster analysis table are the least reliable of the dataset (given the plasticity of the genus) and will be omitted from this analysis. Depending on the statistical test I employ, the traits with more than two values may be converted (ie split into separate binary categories) or omitted. Once the data are tidy I can conduct either Factor Analysis of Mixed Data (FAMD), or, if I convert all variables to binary values, Multiple Correspondence Analysis (MCA). This will cluster like species together and identify the variables that best explain the shape of the data – the traits most responsible for shaping *Apistogramma* phenotype.

Römer U. 2006. Cichlid Atlas Volume 2. 1st Ed. Melle, Germany: Mergus.