New Tools for compositional data analysis of male preference in fish species

Josep A. Martín-Fernández[1] and Michele E.R. Pierotti[2,3]

[1] Departament d'Informàtica i Matemàtica Aplicada, Universitat de Girona, Spain
[2] Department of Biological Sciences, Molecular and Evolutionary Ecology Group, University of Hull, UK
[3] EAWAG Center of Ecology, Evolution and Biogeochemistry, Kastanienbaum, Switzerland

There are few empirical studies of within-population variation in mating preferences and even fewer have examined wild individuals. Here we studied a colour polymorphic population of the Lake Victoria cichlid fish *Neochromis omnicaeruleus* a species in which colour morphs are associated with different sex determining factors. We assessed within-individual consistency and between-individual variation in male mating preferences for the three female colour morphs by testing wild-caught males in three-way choice trials with multiple combinations of different females. Recent and new compositional log-ratio techniques were applied to analyse individual male mating preferences.

Before log-ratio transformation of the data, we modelled observations with zero values, i.e. trials containing zero courtship displays towards one of the stimulus females. Here, zero courtships towards a certain female in a trial might be qualitatively and not only quantitatively different from one or more courtship displays. We modelled these so-called *count* zeros by a new replacement method following a mixed Bayesian-multiplicative approach.

Large individual variation in the strength and direction of male mating preferences for female colour morphs was found by MANOVA applied to log-ratio transformed data. Hierarchical clustering of the compositional data based on Mahalanobis distance revealed the presence of four distinct preference groups corresponding to the three colour morphs in addition to a no-preference class. Most males preferred the rarest morphs in the population, in contrast with previous suggestions proposing assortative mating by genotype-matching between morphs.

Repeatability of male mating preferences of our wild-caught sample was assessed by developing two new multivariate versions of repeatability. One strategy is to replace univariate variances with the corresponding trace of the covariance matrix in the expression of repeatability. Alternatively, a F statistic approach can be followed by applying the relationship between repeatability and F ratio for the analysis of variance.

The stability of all the measures and indices for cluster analysis and repeatability was examined in a bootstrap simulation analysis based on a Monte Carlo process. The resampling procedure consisted of 1000 simulated random samples from the full data set, maintaining the dependence of trials to individual males and/or the groups resulting from the cluster analysis. For each random sample the corresponding indices were calculated and mean, standard deviation, and 2.5% and 97.5% percentiles for the 1000 samples were obtained.