GEA environment,selection and outliers - Melbourne

Bernd Gruber (based on code from Brenna Forrester)

2019-06-07

Contents

Foreword 2

Multivariate GEA: Redundancy Analysis 2

Foreword

This tutorial is based on code from Brenna Forrester, Colorado State University.

The Wolf data set is from Schweizer et al. 2016. Genetic subdivision and candidate genes under selection in North American grey wolves Molecular Ecology 25, 380-402. Dryad doi:10.5061/dryad.c9b25.

Multivariate GEA: Redundancy Analysis

RDA is a multivariate ordination technique that analyzes matricies of loci and environmental predictors simultaneously. RDA determines how groups of loci covary in response to the multivariate environment.

RDA is a two-step analysis in which genetic and environmental data are analyzed using multivariate linear regression, producing a matrix of fitted values. Then PCA of the fitted values is used to produce canonical axes, which are linear combinations of the predictors.

More information on RDA & other multivariate GEAs (such as Random Forest) can be found in our paper: Forester BR, Lasky JR, Wagner HH, Urban DL (2018) Comparing methods for detecting multilocus adaptation with multivariate genotype-environment associations Molecular Ecology 27, 2215-2233.

RDA can be used on both individual and population-based sampling designs. The distinction between the two may not be straightforward in all cases. A simple guideline would be to use an individual-based framework when you have individual coordinates for most of your samples, and the resolution of your environmental data would allow for a sampling of environmental conditions across the site/study area. More on RDA with population level data in the notes below.

Well done!!

That is the end. Well done you finished. Feel free to have another go or have a well deserved beverage!!!