**Instructions for using code**

Below we give instructions for reproducing the analyses used in our publication or applying the method to your own dataset. To use the code, we recommend first downloading all files from the online repository and saving them in the same folder. Note that for all code you must set the working directory to wherever you saved the downloaded data files locally. This is noted in the R code when the command setwd()is called.

1. **Reproduce repertoire size estimates of synthetic datasets used in manuscript**

Running the analyses on the 96 synthetic datasets takes substantial time and uses a lot of memory and processing power. For this reason, we’ve broken the analysis into four separate functions that each analyze a subset of the synthetic data:

* Budgie Clustering small.R : runs analysis on 24 synthetic budgie datasets with 5-15 unique elements
* Budgie Clustering large.R : runs analysis on 24 synthetic budgie datasets with 20-100 unique elements
* LBH Clustering small.R : runs analysis on 24 synthetic long-billed hermit datasets with 5-15 unique elements
* LBH Clustering large.R : runs analysis on 24 synthetic long-billed hermit datasets with 20-100 unique elements

Alternatively, you can just run the function plot synthetic data results.R, which loads the results produced by the functions listed above and creates plots similar to those in the manuscript.

Further reading: Detailed description and instructions for acoustic feature extraction and applying supervised and unsupervised random forest is here:

<https://rpubs.com/marcelo-araya-salas/464516>

<https://rpubs.com/marcelo-araya-salas/475988>

1. **Reproduce acoustic space estimates for synthetic datasets used in manuscript**

Use the function calculate acoustic space.R (note this calls acoustic spaces.R).

Detailed instructions for calculating acoustic space can be found here:

<https://rpubs.com/marcelo-araya-salas/532332>

1. **Reproduce analysis of live bird datasets**

To run the analysis on the field recordings of long-billed hermits, use the file LBH field songs analysis.R .

To run the analysis on the lab recordings of budgerigars, use the file Budgie lab calls analysis.R .

1. **Analyze your own data**

Use the function run analysis on your data.R and follow instructions in the comments.