

Lab 5: strataG

Nora Mitchell

February 2017

Contents

1	Introduction	1
1.1	Installation	1
2	Getting Started	2
3	Using Structure	2
3.1	Running Structure	2
3.2	Evanno et al.	3
3.3	Alternative prior	3
4	References	3

1 Introduction

We've already used Structure (Pritchard et al., 2000), the program that uses multi-locus genotype data and MCMC to perform individual assignment for population genetics analysis.

1. Let's try to use it with a wrapper package in R!
2. *strataG* (Archer et al. 2016) is a toolkit for haploid sequence and multilocus genetic data summaries, and analyses of population structure.
3. Let's look at some assumptions: are populations of equal size (Wang 2016)

1.1 Installation

- Please install Structure v2.3.4 package WITHOUT FRONT END from here http://pritchardlab.stanford.edu/structure_software/release_versions/v2.3.4/html/structure.html
- You will need to know the path of where you have installed the files.

- Add this path to the PATH system variable on your machine. OS specific instructions here: <https://www.java.com/en/download/help/path.xml>
- Install the *strataG* package in R, install the *installr* package in R. Call both libraries using the *library()* function
- Check that your path is listed in your sytem by using the *system.PATH()* function in R (from *installr* library).

2 Getting Started

Let's get our data in shape.

- Call the *adegenet* library, and use the same *read.structure()* code from project 2 (repens.stru) to import the repens data.
- You will now have a “genind” object. We need to convert this to a “gtypes” object for *strataG*
- Use the function *genind2gtypes()* to create a new gtypes object.
- Look at the structure of this object.

3 Using Structure

To actually use structure, use the *structureRun()* function

3.1 Running Structure

- Make sure you use *?structureRun* to look at the function's help page
- Assign function output to a new object called, for instance, “default_output”
- Use a k.range of 2:16 (for sake of time)
- Set num.k.rep equal to 2
- label = “default”
- Use default burnin and numreps

3.2 Evanno et al.

Now let's look at some output

- Use the *evanno()* function to look at Evanno et al. plots for the *default_output*
- Choose value(s) of K to look at
- Make Structure barplots for Ks of interest using the *structurePlot()* function (Hint: look at the structure of your “default_output” object)

3.3 Alternative prior

Let's take Wang (2016)'s advice into account

- To change settings to switch to separate alphas, we need to change the *extraparams* file
- Nora has edited the function and made new function called `runStructureAlt()` in a separate R file
- Run the code to write this function.
- Run this function using same settings as above, name it “alt_output”
- label = “alt”
- Make Evanno plots, look at barplots

Compare outputs, in terms of inferred K

4 References

- Archer, F. I., Adams, P. E. and Schneiders, B. B. (2016), strataG: An R package for manipulating, summarizing and analysing population genetic data. Mol Ecol Resour. doi:10.1111/1755-0998.12559
- Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620.
- Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.
- Wang, J. (2016). The computer program Structure for assigning individuals to populations: easy to use but easier to misuse. Molecular Ecology Resources.