

Pre-workshop instructions

Connor French
connor-french.com

Below you will find all of the necessary software to install for the workshop *Exploring population structure in R with adegenet and sNMF*. Please try to install this before the workshop, and don't hesitate to email me with any problems at french.connor.m@gmail.com. The workshop tutorial is located here, but I would recommend waiting to check it out until the day of the workshop- I'm making some major changes from the last version.

Required and suggested software

R v3.6.x or above. R is a programming language well-suited for data science and has a lot of great packages specifically for working with genetic data. R 4.0.x is the latest version, but some people have been having headaches with it. I'd recommend sticking with R v3.6.x.

RStudio IDE (not necessary, but nice to have). This is a great interactive development environment for coding in R (and can also code in other languages like Python). Think of it as R's version of Jupyter Notebook. Versioning doesn't matter so much with this. If you have never used it before, I'd recommend playing around with it and skim this cheatsheet to familiarize yourself with the layout.

Packages

vcfR. We're using this package to read in the VCF file and convert the file format. It's a nice package for converting among file formats and manipulating VCFs. Run the code below to install the package:

```
install.packages("vcfR")
```

adegenet. We're using this package for running the Discriminant Analysis of Principal Components, but is a nice swiss-army knife genetic analysis package.

```
install.packages("adegenet")
```

LEA. We're using this package for running the sNMF analysis for estimating admixture coefficients. This is the trickiest package to install.

If you have R v4.0.x, use this:

```
if (!requireNamespace("BiocManager", quietly = TRUE))  
  install.packages("BiocManager")
```

```
BiocManager::install("LEA")
```

If you have R v3.6.x, you need to install BiocManager version 3.10. If you have an older version of R and for some reason can't update it, replace `version = "3.10"` with the correct version number from this list.

```
if (!requireNamespace("BiocManager", quietly = TRUE))  
  install.packages("BiocManager")  
BiocManager::install(version = "3.10")  
  
BiocManager::install("LEA")
```

tidyverse. We're using this suite of packages for data manipulation and plotting. The tidyverse is an ecosystem of R packages for data science that make working with tabular data easier.

```
install.packages("tidyverse")
```

rnaturalearth. We're using this package to obtain a basemap for mapping the results.

```
install.packages("rnaturalearth")
```