

SimCo: An R function for calculating similarity coefficients for outputs from Structure.

SimCo is an easy to use function. This document runs through how to use the function in R.

Use requires two steps; (1) importing the Structure outputs and (2) comparing them using the R function.

First you will need to load the R **simco** library.

Windows:

To install simco, the simplest approach is to start R and type:

```
install.packages("simco")
```

This will download the binary from CRAN and install it.

Alternatively, you can download the "simco_1.0.zip" (or the equivalent). Then start R and select (on the menu bar) "Packages" and then "Install package from local zip file...". Find the file "simco_1.0.zip" on your hard drive, and click "Open".

Mac OSX:

Again, the simplest approach is to use:

```
install.packages("simco")
```

This will download the binary from CRAN and install it.

Alternatively, download the compiled version of simco for Mac OS X, a file like "simco_1.0.tar.gz". Then start R and select (on the menu bar) "Packages & Data" -> "Package Installer". Select "Local Binary Package" from the drop-down menu at the top of the window that comes up. Click "Install" at the bottom of the window. Find the package on your drive and click "Open". Finally, close the window.

To load the package, on all platforms you will then need to type:

```
library(simco)
```

Additional help can be found by typing

```
?SimCoImport
```

```
?SimCoef
```

Importing the Structure outputs. Structure outputs are produced in ASCII format and need to be imported into R and have the relevant information extracted from them for analysis. This is handled by the **SimCoImport()** function. It is very important that you don't manipulate the Structure outputs in any way, doing so can destroy the importation process because the function relies on identifying words/syntax within the Structure outputs.

To import the Structure files place the structure files in a directory, use **list.files** to create a list of files to import and then use **SimCoImport** to import the files and append them, including an identifier

In this example I have 3 Structure output files ready for comparison. They are placed in the directory **/Users/orj/Documents/SimCo/structurefiles1/**

Change the working directory of R to that of the Structure output files using **setwd**, then use **list.files** get a list of the files to import (you can use **pattern** to specify a pattern to use if you have other files which you don't want to use in the directory (see **?list.files**)).

```
> setwd("/Users/orj/Documents/SimCo/structurefiles1/")
> myfiles<-list.files()
> myfiles
[1] "K3run_6_f.txt" "K3run_7_f.txt" "K3run_8_f.txt"

> x<-SimCoImport(myfiles)
```

The object “**x**” is now a data frame containing the 3 files, each files being identified with a capital letter. You can check that the importation has worked by typing “**summary(x)**” or “**x**”

Comparing the files.

Now you can use the **SimCoef** function to run a similarity coefficient analysis on the 3 files.

```
> SimCoef(x)
These are the MATRIX permutations:
      [,1] [,2]
[1,] "A"  "B"
[2,] "A"  "C"
[3,] "B"  "C"
```

```
These are the column permutations:
      [,1] [,2] [,3]
[1,]     1     2     3
[2,]     1     3     2
[3,]     2     1     3
[4,]     2     3     1
[5,]     3     1     2
[6,]     3     2     1
```

```
Number of populations (K) = 3
Number of individuals (I) = 74
Number of Structure runs = 3
```

```
Range = 0.975812 - 0.9813705
```

Median Similarity Coefficient = 0.9772389
Mean Similarity Coefficient = 0.9781404
SEM of Similarity Coefficient = 0.001666751

The similarity coefficients were: 0.9813705 0.975812
0.9772389

Summary:

	[,1]	[,2]	[,3]
[1,] A	B	0.981	
[2,] A	C	0.976	
[3,] B	C	0.977	

This output gives you (1) the matrix permutations (i.e. A vs. B, A vs. C etc.), (2) the column order permutations, (3) summary information about the number of populations (K), the number of individuals (I) and the number of Structure runs that were analysed; (4) the range, median, mean and SEM of the similarity coefficients and (5) the similarity coefficients of the (in this case 3) matrix comparisons.

Owen R. Jones (Imperial College, London)
Email: owen.jones@imperial.ac.uk