SwingerV0.92 User Manual

J. Sandoval-Castillo; C. R. M. Attard; S. Marri; C. Brauer; L. M. Möller; L. B. Beheregaray
2016

Swinger is a workbench to select the best groups of breeders for conservation breeding programs without pedigree information.

One of the main goals of genetic conservation management is to preserve the genetic variability of wild populations to prevent the extinction of endangered species. Unfortunately, most endangered species have low numbers of reproductive populations of very small sizes. This is commonly reflected in low levels of genetic variation that make *in situ* conservation efforts inefficient. In most cases, breeding of captive individuals is the best *ex situ* conservation alternative.

Captive breeding programs aim to retain most of the natural remaining variation and to transmit it to the future generations. The best strategy to maintain genetic diversity is minimizing the kinship between reproductive pairs. Nevertheless, for a great number of endangered species there is no accurate pedigree information, so precise determination of kinship is impossible. Pairwise relatedness estimators were not designed to determine kinship between pairs of individuals and consequently have inherent limitations on pedigree constructions. However, when there is no knowledge about the pedigree, or pedigree construction is not practical, the use of genetic markers to estimate relatedness and infer kinship is often the only alternative.

Swinger uses pairwise relatedness to determine the best groups of breeders to effectively preserve most genetic variation in captive breeding programs. Although most captive breeding programs use a limited number of breeders, the increment of possible pairwise combinations is a factorial function of the number of individuals. In other words, in polygamous reproductive systems tens of individuals can have millions of possible groups. Identifying the best possible combination of groups is not just a tedious and time consuming process, but also prone to human error. Swinger implements an automatic algorithm to determine the best possible combination of groups based on thresholds of individual internal relatedness, pairwise relatedness and average relatedness in groups.

SYSTEM REQUIREMENTS

Swinger is a Perl script with a java graphic interface, therefore can be run in Windows, Linux or Unix. It has been tested in Ubuntu 14.04.1, Windows 10 and Mac OS X 10.11. Swinger can be run on any standard desktop computer, although depending on the size of data set can require several GB of RAM and is not multi-threaded.

Java 8 RE and Perl 5 need to be installed. The Perl script also depends on the modules "Math::Combinatorics", "Array::Utils" and "MCE".

Installing SwingerV0.91

All necessary instructions and links are available on

http://www.molecularecology.flinders.edu.au/molecular-ecology-lab/software/swinger/swinger/.

Install Perl 5 or newer

Windows

Download an installer for Perl from http://www.perl.org/get.html and run the installer.

Linux and Mac

Almost all Unix and Mac systems have Perl already installed. In a command line or terminal type "perl – v" to find out which version.

Install Java 8 or newer

Windows

Download an installer for Java 8 or newer from http://www.oracle.com/technetwork/java/javase/downloads/index.htmljava and run the installer.

Ubuntu

Execute the following command as root user

```
"apt-get install openjdk-8-jre"
```

Mac

Apple supplies their own version of Java on their computers.

Install Perl modules

The modules can be downloaded from http://search.cpan.org/~allenday/Math-Combinatorics-0.09/lib/Math/Combinatorics.pm and http://search.cpan.org/dist/Array-Utils/Utils.pm.

However, for an easy download and installation of Perl modules, we recommend the installation of ActivePerl or cpan.

To install ActivePerl simply download and run the suitable installer from http://www.activestate.com/activeperl/downloads. To install the modules just start ActivePerl and look for the module required.

To install cpan just type the following command in a command line or terminal

```
"cpan App::cpanminus"
```

For the three modules that need to be installed, this will be

```
"cpan Math::Combinatorics"
```

```
"cpan Array::Utils"
```

```
"cpan MCE"
```

HOW TO USE SWINGER GUI

Input files

Swinger uses inbreeding and pairwise relatedness estimations from molecular markers to determine the best combination of reproductive groups. The main input file is a square matrix of the pairwise relatedness values with a measure of inbreeding in a tab-delimited text file.

On the first row are the ID of the samples, the second and subsequent rows are the information per sample as follows: first column = sex (M = male or F = female), second column = internal relatedness, third column = ID, and fourth column onwards = corresponding pairwise relatedness estimates. Swinger requires the rows to be ordered so the female individuals are first.

			Sample1	Sample2	Sample3	Sample4
F	0.002	Sample1	0	-0.01	-0.12	0.01
F	0.342	Sample2	-0.01	0	-0.02	0.02
M	0.001	Sample3	-0.12	-0.02	0	-0.001
M	-0.025	Sample4	0.01	0.02	-0.001	0

Example of input file in a matrix format

If you don't need or have internal relatedness estimates, you can put these values as equal to 0 and select 1 as your threshold for internal relatedness (see section "Internal relatedness").

Since some software used to calculate relatedness produce a table as output, we include the option to convert a table to matrix format. This function requires two input files:

1. A table of internal relatedness in a tab-delimited text file, where each row has the information for one sample as follows: first column = ID, second column = sex (M = male or F = female), and third column = internal relatedness value. Swinger requires the rows to be ordered so the female individuals are first.

Sample1	F	0.002
Sample2	F	0.342
Sample3	M	0.001
Sample4	M	-0.025

Example of internal relatedness input file in a table format

Again, if you don't need or have internal relatedness, you can put these values as equal to 0 and select 1 as your threshold for internal relatedness.

2. A table of pairwise relatedness in a tab-delimited text file, where first column = first ID, second column = second ID, third column = relatedness value.

Sample1	Sample2	-0.01
Sample2	Sample3	-0.02
Sample3	Sample4	-0.001
Sample1	Sample3	-0.12
Sample2	Sample4	0.02
Sample1	Sample4	0.01

Example of pairwise relatedness input file in a table format

RUNNING SWINGER

Execute Swinger GUI

Double click on the java executable "`SwingerGUI.jar`".

Alternatively, in the command line within the folder containing SwingerGUI.jar execute the command "`java –Xmx1024m –Xms512M –jar SwingerGUI.jar`" to start the program.

Load input files

Once the input files are built, you must load them into Swinger. They must be placed in a path that has no spaces. To load the input files you have several options:

Pairwise relatedness table

    Activating the menu File | Open PRT

or

    Clicking on the "Open Pairwise Relatedness Table" button and selecting the corresponding file

or

Writing the path to the corresponding file in the text field left of the "Open Pairwise Relatedness Table" button.

or

Using the short cut "Alt+1", and selecting the corresponding file

Internal relatedness table

Activating the menu File | Open IRT

or

Clicking on the "Open Internal Relatedness Table" button and select the corresponding file

or

Writing the path to the corresponding file in the text field left of the "Open Internal Relatedness

Table" button

or

Using the short cut "Alt+2", and selecting the corresponding file

Pairwise relatedness matrix

Activating the menu File | Open RM

or

Clicking on the "Open Relatedness Matrix" button and selecting the corresponding file

or

Writing the path to the corresponding file in the text field left of the "Open Relatedness Matrix" button

or

Using the short cut "Alt+I", and selecting the corresponding file

or

Activating the menu File | Recent Matrices | Matrix file

Select output files

To select the output files to be saved, you also have several options:

Converted pairwise relatedness matrix

This is the pairwise relatedness matrix with internal relatedness values created by Swinger when the input is supplied as tables.

Activating the menu File | Save RM and selecting the desired path and file name

or

Clicking on the "Save Relatedness Matrix" button and selecting the desired path and file name

or

Writing the desired path and file name in the text field left of the "Save Relatedness Matrix" button

or

Using the short cut "Alt+3" and selecting the desired path and file name

Best combination of groups

This is the best combination of groups for breeding found by Swinger according to the user-defined parameters (see "Parameters" below).

Activating the menu File | Save BC and selecting the desired path and file name

or

Clicking on the "Save Best Combination" button and selecting the desired path and file name

or

Writing the desired path and file name in the text field left of the "Save Best Combination" button

or

Using the short cut "Alt+O" and selecting the desired path and file name

Select parameters

Once all the input files are loaded and the output files indicated, you have to select the setup and thresholds for the breeding groups. All parameters require a value unless otherwise indicated below, in which case they should be left blank.

Combination of groups set up: These parameters will depend on the reproductive system of you species (monogamy, polygamy, polygyny, polyandry, etc) and the number of potential breeders.

> #Groups? (NG) = Number of groups to be created. No individuals will be used more than once within or among the groups.

> #Females?(NF) = Number of females per group.

> #Males? (NM)= Number of males per group.

Combination thresholds: It is important to try strict thresholds to get the best possible groups. The values will depend on the reproductive system of your species and the levels of relatedness in your data.

> Max Internal Relatedness (IR) = The maximum relatedness value allowed in each individual to be considered for the formation of groups. This is required when Max IR Males and Max IR Females are not indicated and you want the same values for males and females.

> Max Pairwise Relatedness (PR) = The maximum relatedness value allowed between each pair of individuals to be considered for the formation of groups. In groups, when you want different values for pairwise interaction depending on gender (male-male, male-female, female-female), this value represents the maximum relatedness allowed between straight couples (female-male).

> Max Aver Relatedness per Group (ARG) = The maximum average relatedness allowed in a group. This is only required when Max Aver Relatedness Couples per Group is not indicated.

> Max Aver Relatedness per Combination (ARC) = The maximum average group relatedness per combinations. This is only required when Max Aver Relatedness Couples per Group is not indicated.

In many cases the best set of groups will be composed of individuals with low relatedness values. However, in some cases the reproductive system (e.g. monogamy) or high relatedness in a particular gender makes it more practical to have independent thresholds per gender and per possible pairs (male-male, male-female, female-female). Because of this, we have included the following optional thresholds.

> Max IR Males (IRM) = The maximum internal relatedness value allowed in each male to be considered for the groups. This is only required when Max Internal Relatedness is not indicated and you want different values for males and females.

> Max IR Females (IRF) = The maximum internal relatedness value allowed in each female to be considered for the groups. This is only required when Max Internal Relatedness is not indicated and you want the different values for males and females.

> Max PR Males (PRM) = The maximum relatedness value allowed between each pair of males to be considered for the groups. This is only required when you want different values for pairwise interaction depending on gender (male-male, male-female, female-female).
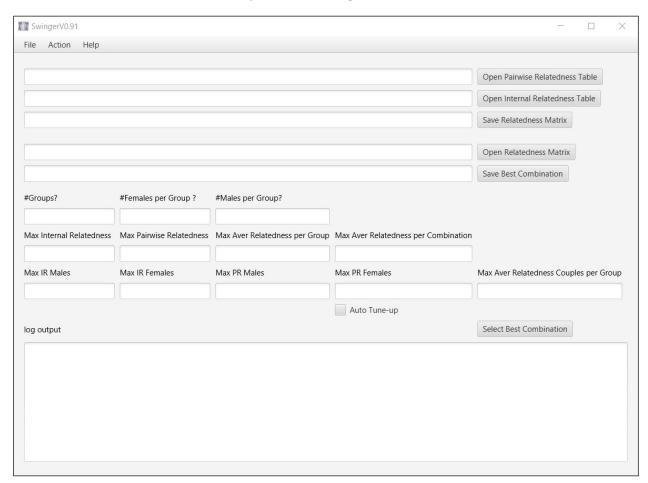
> Max PR Females (PRF) = the maximum relatedness value allowed between each pair of females to be considered for the groups. This is only required when you want different values for pairwise interaction depending on gender (male-male, male-female, female-female).

Max Aver Relatedness Couples per Group (ARCG) = The maximum average relatedness in the couples (male-female) of a group. This is required when Max Aver Relatedness per Group is not indicated and you want different values for pairwise interaction depending on gender (male-male, male-female, female-female).

If you don't have any idea of the possible best values, for initial runs we recommend using the average relatedness of your data as your PR, your ARF 10-50% lower than your PR, and your ARB 10-50% lower than your ARP. The log output printed in the GUI when a run finishes will indicate how many parties are possible, the number that passed your relatedness thresholds and the number of solutions with the values selected. Based on this, the values may be altered and Swinger re-run.

We have also added an automatic tune-up function. To activate, tick the checkbox "Auto Tune-up" before running the software. Swinger will then play with the user-defined thresholds for pairwise relatedness and average pairwise relatedness within groups. If they are too strict, with no sets of pairs or groups passing the thresholds, Swinger will increment upwards by 2% the values of these thresholds and then automatically re-run. If they are too flexible, with possible combinations of groups more than three, Swinger will instead reduce these thresholds by 10% and re-run.

Although we recommend the use of the tune-up function, you still need relatively good initial thresholds otherwise this function will be extremely time consuming.

Swinger graphic interface with dialog boxes to select input and output files, and parameters.

Start algorithm

How to start the algorithm for finding the best groups depends on whether the input file format was tables or matrices, as described below. Note that the GUI will appear non-responsive from when the algorithm starts until it finishes.

Tables

Once the internal and pairwise relatedness tables are loaded, the matrix and best combination of groups output files are indicated, and all the parameters are set up, you can execute the algorithm:

> Activating the menu Action | Convert T2M

or

> Clicking on the "Select Best Combination" button

or

> Using the short cut "Alt+C"

Matrices

Once the pairwise relatedness matrix is loaded, the best combination of groups output file is indicated, and all the parameters are set up, you can execute the algorithm:

> Activating the menu Action | Select BFG

or

> Clicking on the "Select Best Combination" button

or

> Using the short cut "Alt+S"

OUTPUT FILES

Swinger will create a text file with a list of up to three solutions of *x* number of groups (#Parties) that pass the thresholds set up. For each combination of groups, an individual will only appear up to once. The first line will show the combination number and the second line will show two values of average relatedness for that combination (between all the individuals and between couples of different sex). Then, for each group within the combination there will be a line indicating the group number followed by the pairwise relatedness matrix. The matrix contains the sample names, as they appear in your input files, at the first row and column. The first cell of each matrix will show two values of average pairwise relatedness in the group (between all the individual and between couples of different sex).

Average relatedness per group within combination 1

Between all individuals / between pairs of different sex

Average pairwise relatedness within group 1 of combination 1

Between all individuals / between pairs of different sex

```
Selected family group number 1:
Average relatedness for all 2 families = -0.2228/-0.2475

Family 1 in group 1
-0.2520/-0.2838 SPBr01  SPBr39  SPBr15  SPBr65
SPBr01   0     -0.166  -0.204  -0.355
SPBr39  -0.166   0     -0.44   -0.136
SPBr15  -0.204  -0.44    0     -0.211
SPBr65  -0.355  -0.136  -0.211   0
1 15 35 63

Family 2 in group 1
-0.1935/-0.2112 SPBr31  SPBr54  SPBr16  SPBr35
SPBr31   0     -0.148  -0.295  -0.135
SPBr54  -0.148   0     -0.184  -0.231
SPBr16  -0.295  -0.184   0     -0.168
SPBr35  -0.135  -0.231  -0.168   0
10 21 36 48


Selected family group number 2:
Average relatedness for all 2 families = -0.2231/-0.2241

Family 1 in group 2
-0.1935/-0.2112 SPBr31  SPBr54  SPBr16  SPBr35
SPBr31   0     -0.148  -0.295  -0.135
SPBr54  -0.148   0     -0.184  -0.231
SPBr16  -0.295  -0.184   0     -0.168
SPBr35  -0.135  -0.231  -0.168   0
10 21 36 48

Family 2 in group 2
-0.2527/-0.2370 SPBr01  SPBr46  SPBr15  SPBr48
SPBr01   0     -0.298  -0.204  -0.262
SPBr46  -0.298   0     -0.306  -0.176
SPBr15  -0.204  -0.306   0     -0.27
SPBr48  -0.262  -0.176  -0.27    0
1 18 35 53
```

Example of output file with the best two combination, each consisting of two groups