

Friends and Family version 21 README

Last updated on 02 March 2021

Table of Contents

Friends and Family version 21 README	1
License.....	2
Starting Friends and Family.....	2
Input files	2
Table input file format	2
Tri-matrix input file format	2
Running Friends and Family.....	2
Description of how Friends and Family performs the analysis.....	3
Output files	3
Known issues and troubleshooting.....	3
Citation.....	4
References	5

License

Use of Friends and Family is permitted as per the GNU GENERAL PUBLIC LICENSE Version 3, 29 June 2007.

Starting Friends and Family

Friends and Family can be downloaded from here: <https://github.com/DeondeJager/Friends-and-Family>.

It may be possible that the program has a file extension of *.programme after unzipping the folder. In this case, the user should manually change the file extension to .exe to make the file executable. It may happen that the anti-virus software installed on the computer scans the executable file for malware before it allows the program to open. It should not find any malware and should allow Friends and Family to run thereafter.

Input files

Friends and Family requires as input either a tri-matrix or table of pairwise relatedness values in comma-separated values (CSV) format (without UTF-8 encoding). The table input file can be generated in a program such as [Coancestry](#)¹ and the tri-matrix input file can be generated in a program such as [GenAlEx](#)².

Table input file format

Sample names must be in the first and second columns, with the pairwise relatedness values for each comparison in the third column.

Tri-matrix input file format

Sample names must be in the first column and first row only. Pairwise relatedness values of a sample compared to itself must be equal to 0 (not 1), which is the standard output of the pairwise relatedness tri-matrix generated by GenAlEx.

See the example input files for the format.

Please see the **Known issues and troubleshooting** section below for more details on input file formatting if there are problems reading the file or running the program.

Running Friends and Family

Once the input file has been loaded, select either the `Tri-matrix` or the `Single column` options in the `Data format` box for tri-matrix and table input files, respectively. Thereafter, the user must specify the names of the output files and the location where they will be saved (see **Output files** below for more detail on the output files generated). Once this has been done, the user must enter their chosen relatedness cut-off value in the `Relatedness cut-off` box. The default value is 0.25, but can be any number between -1 and 1. The user also has the option to `Produce binary output files`, but this is not necessary for the program to perform the analysis (see **Output files** for a description of the binary output files).

Description of how Friends and Family performs the analysis

The program selects the individual with the least number of relatives in the data set as the reference individual for that round. It then removes all the individuals who are related to the reference individual (i.e. have a pairwise relatedness value equal to or greater than the given cut-off value). This reference individual is then labelled as a "friend" and is retained in the data set. In the second round, the next sample with the least number of relatives is chosen (using the data set generated by round 1, where individuals have been removed) and again removes all relatives of this new reference individual, who is then also labelled as a "friend". This process continues until there are no relatives remaining in the data set, so that a data set is generated where all samples are unrelated to each other.

Output files

The three standard output files produced are a `friends.csv` output file, with the list of unrelated individuals, a `family.csv` output file, with the identity of the reference sample for each round and its family members that were removed during that round, and a file called `wa3b2.csv`, which is generated by the program as a step during analysis that contains a square matrix of the pairwise relatedness input data. It is important to note that the individuals shown in the `family.csv` output file may not be all the family members of that reference individual, since some of its family members may have been removed from the data set during previous rounds.

Furthermore, the user can select the option to output binary files. These CSV files show, for each round, the status of each sample (friend or not), the number of relatives each sample has and their identities, as well as the average relatedness values of the samples with the fewest number of relatives if this calculation was required (i.e. if two samples have the same number of relatives, the sample with the lowest average relatedness across its relatives is used as the reference sample for that round).

The example output files show what these files should typically contain.

Known issues and troubleshooting

Friends and Family v21 has been tested on Windows 10, 8 and 7 and works on all three operating systems. It should also work on older versions of Windows, but has not been tested on these operating systems.

- The input file should be reloaded each time an analysis is performed, otherwise an error box reading "The number of data points (a number) do not fit a tri-matrix" appears. Clicking OK in this error box closes the program, after which the user has to restart the program to perform the analysis.
- Occasionally, different versions of Excel save CSV files with delimiters that are somehow not commas and thus the input file is not recognised or readable by Friends and Family. In this case, an error message will appear, or the program will produce empty `friends.csv` and `family.csv` output files. The user should attempt to edit the input file in another program, like a text editor such as Notepad++, and replace the delimiters with commas and then save the input file as a CSV file again. The reason Excel saves the CSV files with different delimiters (such as a

semi-colon, instead of a comma), is when the user's system uses a comma as the standard decimal separator. This means the relatedness values in the input file contains commas as the decimal separator and the field delimiter used is a semi-colon instead of a comma. To check your system default decimal separator, go to Control Panel --> Region --> Additional Settings.

- Error: "Invalid input data: Row 2". This error seems to appear when the end-of-line (EOL) character is not recognized by Friends and Family. It seems that the program does not like CSV input files generated on Mac, which by default uses the "CSV UTF-8 (Comma delimited) (*.csv)" option when saving CSV files from Excel, which uses the "CR" EOL character. I use [Notepad++](#) to show EOL characters: open the CSV file in Notepad++, then select View --> Show Symbol --> Show All Characters. The simplest solution is to save CSV files from Excel using the "CSV (Comma delimited) (*.csv)" option, which will then use the EOL character "CRLF" (the DOS default), which Friends and Family recognises.
Another option to solve this error is to convert the CSV input file generated on a Mac to DOS or Unix format using the [dos2unix](#) utility. Friends and Family seems to recognise the Unix EOL character "LF" just fine, so converting from Mac to Unix should do the trick ([mac2unix](#)), but one could go a step further and convert the Unix format to DOS ([unix2dos](#); there is no [mac2dos](#) utility). Thanks to Karen Holm for bringing this error to my attention.
- Error: "Invalid relatedness cut off value '0.25' is not a valid floating point value". This error occurred with version 20 of Friends and Family, but has been fixed in version 21 onwards. Always use a full stop "." in the program (in the cut-off value box), as it will correct internally if your system default decimal separator is a comma.
The error arose because of the default decimal separator on the user's system. The program expects a full stop ".", but when the default separator is a comma ",", then the program does not recognise the value entered into the cut-off value box as a valid number. Changing the separator in the cut-off value box to a comma results in the program not recognising the data values in the input file as valid numbers (as these contain a full stop as a decimal separator). To check your system default decimal separator, go to Control Panel --> Region --> Additional Settings. You do not have to change the default separator on your system. You merely have to edit the decimal separator and field delimiter in your input file to a full-stop and comma, respectively, in a text editor such as [Notepad++](#) (using `ctrl+h` to search and replace) and save the file as CSV.

For additional error reporting or if you have solved a known or new error and would like to share it, please contact Deon de Jager at dejager4@gmail.com.

Citation

de Jager D, Swarts P, Harper C, Bloomer P (2017) Friends and Family: A software program for identification of unrelated individuals from molecular marker data. *Molecular Ecology Resources*, e225-e233, doi:10.1111/1755-0998.12691. <https://onlinelibrary.wiley.com/doi/full/10.1111/1755-0998.12691>

References

- 1 Wang, J. Coancestry: A program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Mol. Ecol. Resour.* **11**, 141-145, doi:10.1111/j.1755-0998.2010.02885.x (2011).
- 2 Peakall, R. & Smouse, P. E. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**, 2537-2539, doi:10.1093/bioinformatics/bts460 (2012).